

# Evaluierung des Deutschen Satellitendatenarchivs auf Nachhaltigkeit und Effizienz

Evaluation of the German Satellite Data Archive in Terms of  
Sustainability and Efficiency

## Masterarbeit

im Fachgebiet Systems Engineering



vorgelegt von: Max Wegner

Fakultät: Elektrotechnik und Informationstechnik

Matrikelnummer: 05 06 04 09

Bearbeitungsbeginn: 01.07.2012

Abgabetermin: 02.04.2013

Laufende Nummer: 271

© 2012

Dieses Werk einschließlich seiner Teile ist **urheberrechtlich geschützt**. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Autors unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen sowie die Einspeicherung und Verarbeitung in elektronischen Systemen.

# Evaluierung des Deutschen Satellitendatenarchivs auf Nachhaltigkeit und Effizienz

Evaluation of the German Satellite Data Archive in Terms of  
Sustainability and Efficiency

## Masterarbeit

im Fachgebiet Systems Engineering



vorgelegt von: Max Wegner

Fakultät: Elektrotechnik und Informationstechnik

Matrikelnummer: 05 06 04 09

Erstgutachter: Prof. Dr. Werner Kohl

Zweitgutachter: Dipl.-Inf. Stephan Kiemle

© 2012

Dieses Werk einschließlich seiner Teile ist **urheberrechtlich geschützt**. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtgesetzes ist ohne Zustimmung des Autors unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen sowie die Einspeicherung und Verarbeitung in elektronischen Systemen.

---

## Eidesstattliche Erklärung

---

Name

---

Vorname

1. Ich erkläre hiermit, dass ich die vorliegende Masterarbeit selbständig verfasst und noch nicht anderweitig zu Prüfungszwecken vorgelegt habe.

Sämtliche benutzte Quellen und Hilfsmittel sind angegeben, wörtliche und sinngemäße Zitate sind als solche gekennzeichnet.

München, den 2. April 2013

---

MAX WEGNER

2. Ich erkläre mein Einverständnis, dass die von mir erstellte Masterarbeit in die Bibliothek der Hochschule München eingestellt wird. Ich wurde darauf hingewiesen, dass die Hochschule in keiner Weise für die missbräuchliche Verwendung von Inhalten durch Dritte infolge der Lektüre der Arbeit haftet. Insbesondere ist mir bewusst, daß ich für die Anmeldung von Patenten, Warenzeichen oder Geschmacksmuster selbst verantwortlich bin und daraus resultierende Ansprüche selbst verfolgen muß.

München, den 2. April 2013

---

MAX WEGNER

---

## Zusammenfassung

Das Deutsche Satellitendatenarchiv (DSDA) ist eine Kernkomponente des Erdbeobachtungszentrums (*Earth Observation Center*, EOC) des Deutschen Zentrums für Luft- und Raumfahrt (DLR). Diese Arbeit evaluiert das DSDA anhand der Long Term Data Preservation Common Guidelines der Europäischen Raumfahrtagentur ESA bezüglich seiner Nachhaltigkeit und Effizienz. Auf dieser Grundlage werden zwei Maßnahmen für eine gesteigerte Effizienz, Wirtschaftlichkeit beziehungsweise Nachhaltigkeit diskutiert und näher betrachtet werden. Diese sind die Zusammenlegung der Archive in Oberpfaffenhofen und Neustrelitz sowie die Implementierung eines Produkttyps für Dokumente in der Product Library.

## Abstract

The German Satellite Data Archive (DSDA) is a key component within the Earth Observation Center (EOC) which belongs to the German Aerospace Center (DLR). This thesis evaluates the DSDA on behalf of the Long Term Data Preservation Common Guidelines of the European Space Agency ESA regarding its sustainability and efficiency. On this basis, two measures to improve the economical or technical efficiency or the sustainability will be discussed further. These are the combination of the two separate archives in Oberpfaffenhofen and Neustrelitz and the implementation of a new product type for documents within the Product Library.

## Inhaltsverzeichnis

<b>Abkürzungsverzeichnis</b>	<b>III</b>
<b>Abbildungsverzeichnis</b>	<b>V</b>
<b>Tabellenverzeichnis</b>	<b>VI</b>
<b>1. Einleitung</b>	<b>1</b>
1.1. Problemstellung und Ziel dieser Masterarbeit . . . . .	1
1.2. Vorgehensweise . . . . .	1
<b>2. Das Deutsche Satellitendatenarchiv</b>	<b>3</b>
2.1. Missionen und Daten . . . . .	3
2.2. Daten- und Informationsmanagementsystem . . . . .	4
2.2.1. Product Library . . . . .	5
2.2.2. Archiv . . . . .	6
2.2.2.1. HSM . . . . .	6
2.2.2.2. Bänder versus Festplatten . . . . .	7
2.2.2.3. Archivwartung . . . . .	9
2.3. Standorte Oberpfaffenhofen und Neustrelitz . . . . .	10
2.4. Systemgrenzen . . . . .	10
2.4.1. Beschreibung des Umfeldes . . . . .	10
2.4.2. Blackbox- und Greybox-Betrachtung . . . . .	11
2.4.3. Systemabgrenzung . . . . .	11
<b>3. Long Term Data Preservation</b>	<b>13</b>
3.1. Einführung . . . . .	13
3.2. Bearbeitung . . . . .	13
3.3. Auswertung . . . . .	14
3.3.1. Preserved Data Set Composition . . . . .	14
3.3.2. Archive Operation and Organization . . . . .	16
3.3.3. Archive Security . . . . .	18
3.3.4. Data Ingestion . . . . .	19
3.3.5. Archive Maintenance . . . . .	20
3.3.6. Data Access and Interoperability . . . . .	22

3.3.7. Data Exploitation and Re-Processing . . . . .	23
3.3.8. Data Appraisal and Purge Prevention . . . . .	25
3.4. Bewertung . . . . .	25
3.4.1. Auswahl der zu behandelnden Themen . . . . .	26
<b>4. Zusammenlegung der Archive in Oberpfaffenhofen und Neustrelitz</b>	<b>28</b>
4.1. Voraussetzungen . . . . .	28
4.2. Vorteile . . . . .	29
4.3. Mögliche Lösungsansätze . . . . .	30
4.3.1. Nutzung des DIMS-Services UIF . . . . .	30
4.3.2. Verwendung eines Cluster-Dateisystems . . . . .	31
4.3.3. Offline-Transport der Daten . . . . .	31
4.4. Netzwerkanbindung und Kosten/Nutzen-Betrachtung . . . . .	32
4.5. Schlussfolgerung . . . . .	33
<b>5. Speicherung von Missionsdokumenten in der Product Library</b>	<b>34</b>
5.1. Emulation, Migration, Description . . . . .	34
5.2. Geeigneter Speicherort . . . . .	35
5.2.1. Dedizierter Missionsserver . . . . .	35
5.2.2. Zentrales Dokumentenmanagement . . . . .	35
5.2.3. Speicherung direkt in der Product Library . . . . .	35
5.2.3.1. Konfiguration eines neuen Dokumenten-Produkttyps	36
5.2.4. Offene Punkte . . . . .	36
<b>6. Fazit und kritische Bewertung</b>	<b>39</b>
<b>Literaturverzeichnis</b>	<b>40</b>
<b>A. Anhang</b>	<b>i</b>
A.1. Produktmodell MetaItemFile . . . . .	i
A.2. LTDP Common Guidelines . . . . .	vi

## Abkürzungsverzeichnis

DFD .....	Deutsches Fernerkundungsdatenzentrum; Institut des DLR
DFN .....	Deutsche Forschungsnetz
DIMS .....	Data and Information Management System; ein im EOC entwickeltes Software-Framework für den Betrieb des Nutzlastbodensegments von Erdbeobachtungsmissionen
DLR .....	Deutsches Zentrum für Luft- und Raumfahrt e.V.
DSDA .....	Deutsches Satellitendatenarchiv; das Langzeitarchiv des EOC
EOC .....	Earth Observation Center; bestehend aus den Instituten DFD und IMF
EOWEB .....	Earth Observation on the Web; Nutzerschnittstelle zum Langzeitarchiv, <a href="http://eoweb.dlr.de">http://eoweb.dlr.de</a>
ESA .....	European Space Agency; Europäische Weltraumagentur
FTP .....	File Transfer Protocol
GSOC .....	German Space Operations Center (Deutsches Raumfahrtkontrollzentrum)
HDF .....	Hierarchical Data Format; Dateiformat zur Messdatenspeicherung, Verwendung z.B. in der Atmosphärenforschung
HSM .....	Hierarchical Storage Management
IIF .....	Item Information File; eine die Archivdaten beschreibende Metadaten-XML-Datei
IMF .....	Institut für Methodik der Fernerkundung
JPEG .....	Joint Photographic Experts Group; verlustbehaftetes Bildkompressionsverfahren
LTDP .....	Long Term Data Preservation; Arbeitsgruppe der ESA
NetCDF .....	Network Common Data Format; wissenschaftliches Datenformat, Verwendung z.B. in der Geoinformatik
OAIS .....	Open Archival Information System; Referenzmodell für Langzeitarchive
OGC .....	Open Geospatial Consortium
PL .....	Product Library

PNG .....	Portable Network Graphics; verlustfreies Bildkompressionsverfahren
RAID .....	Redundant Array of Independent Disks
SAFE .....	Standard Archive Format for Europe; wissenschaftliches Container-Dateiformat, von der ESA spezifiziert
SAN .....	Storage Area Network; verteilte Speicherstrukturen in dedizierten Netzwerken
TIFF .....	Tagged Image File Format; zumeist verlustfreies Dateiformat für Grafiken, wenig komprimiert
UIF .....	User Information Services Interface Framework
XML .....	Extensible Markup Language



## Abbildungsverzeichnis

2.1. DIMS-System-Übersicht . . . . .	5
2.2. HSM-Übersicht . . . . .	7
2.3. Robot Library beim EOC. Links sind die Laufwerke zu sehen, rechts ein Ausschnitt auf die Bandstellplätze, im Vordergrund der Roboterarm.	8
5.1. Konfiguration des neuen Bereichs "DemoDocuSpace" . . . . .	37
5.2. Definition des neuen Produkttyps "ProductTypeDocumentation" . .	38
5.3. Ansicht des neuen Produkttyps im Operating Tool . . . . .	38

## **Tabellenverzeichnis**

3.1. Grad der Erfüllung für die LTDP-Guideline-Kategorien . . . . .	26
3.2. Grad der Erfüllung der LTDP-Guideline-Themenbereiche . . . . .	26
4.1. Mehrkosten für die Netzwerkanbindung . . . . .	32

# 1. Einleitung

## 1.1. Problemstellung und Ziel dieser Masterarbeit

Das Deutsche Satellitendatenarchiv (DSDA) ist eine Kernkomponente des Erdbeobachtungszentrums (*Earth Observation Center*, EOC) des Deutschen Zentrums für Luft- und Raumfahrt (DLR).

Diese Arbeit soll das DSDA anhand einer geeigneten Vorlage bezüglich seiner Nachhaltigkeit und Effizienz evaluieren. Auf dieser Grundlage sollen zwei mögliche Maßnahmen für eine gesteigerte Effizienz, Wirtschaftlichkeit beziehungsweise Nachhaltigkeit diskutiert und näher betrachtet werden.

## 1.2. Vorgehensweise

**Kapitel 2** beschreibt zunächst das Deutsche Satellitendatenarchiv, seine Ziele und die gespeicherten Daten. Weiterhin wird ein genauer Blick auf das Daten- und Informationsmanagementsystem inklusiver der Product Library und dem eigentlichen Datenarchiv geworfen. Abschließend geht dieses Kapitel noch näher auf die beiden Standorte Oberpfaffenhofen und Neustrelitz, sowie die Systemgrenzen ein, wie sie in dieser Arbeit betrachtet werden sollen.

**Kapitel 3** erläutert, um was es sich bei der Long Term Data Preservation Working Group (LTDP) der Europäischen Weltraumagentur handelt und was die LTDP-Common-Guidelines beinhalten. Anschließend wird das DSDA anhand dieser Richtlinien bezüglich seines Erfüllungsgrades bewertet. Das Ergebnis dieser Bewertung wird noch einmal reflektiert, um zuletzt zwei Themenbereiche heraus zu greifen, die in den folgenden Kapiteln näher betrachtet werden sollen.

**Kapitel 4** untersucht die Möglichkeit, die Archive in Oberpfaffenhofen und Neustrelitz zusammenzulegen. Hierfür werden nach der Untersuchung der Voraussetzungen die sich daraus ergebenden Vorteile erläutert sowie drei Lösungsansätze beschrieben.

**Kapitel 5** führt eine Konfiguration vor, die es erlaubt, innerhalb der Product Library neben den Datenprodukten ebenso die Dokumentation für eine Satellitenmission abzuspeichern. Auf diese Weise wären die Daten und die sie beschreibenden Dokumente im selben Archiv untergebracht, was viele Jahre nach einem Missionsende das Auffinden dieser Dokumente erheblich erleichtern würde.

**Kapitel 6** resümiert zum Schluss kritisch die erzielten Ergebnisse und gibt einen Ausblick auf weitere Aktivitäten in der Zukunft.

## 2. Das Deutsche Satellitendatenarchiv

Satellitengestützte Fernerkundung ist heutzutage in nahezu allen Lebensbereichen unerlässlich. Mit ihrer sowohl weltweiten wie regionalen Abdeckung in unterschiedlichen Spektralbereichen und Auflösungen deckt sie die vielfältigsten Anwendungsgebiete ab. Hierzu gehören Landnutzungskarten, Landoberflächen- und Meerestemperaturen, die Konzentration atmosphärischer Spurengase, die Stärke und Ausdehnung von Wolken, digitale Geländemodelle sowie optische und Radaraufnahmen. Darauf basierende Daten sind sowohl für wissenschaftliche, staatliche und zunehmend auch privatwirtschaftliche Anwendungen von hohem Interesse. Neben der räumlichen Abdeckung ist ebenso die Zeitliche zunehmend gefragt. Diese Möglichkeit ergibt sich aus dem vorhandenen Datenbestand der letzten Jahrzehnte. Somit ist die Betrachtung von Änderungen auf der Erdoberfläche, sei es durch menschlichen oder natürlichen Einfluss, über einen längeren Zeitraum hinweg möglich.<sup>1</sup>

Zur Speicherung, Verarbeitung und Nutzbarmachung solcher Datenmengen ist eine gut funktionierende und zuverlässige informationstechnische Infrastruktur nötig. Diese Infrastruktur muss auf der einen Seite leicht bedien- und wartbar sein, auf der anderen Seite jedoch auch so flexibel, dass sie mit den vielfältigen Datenprodukten nicht überfordert ist. Zudem müssen die Daten über einen langen Zeitraum hinweg sicher archiviert und wieder abrufbar sein. Dieser Zeitraum geht bei weitem über die Lebensdauer des jeweiligen Satelliten hinaus. Besonders für Zeitreihen müssen zumindest die Rohdaten vom Satelliten möglichst "für immer" aufgehoben werden, da es keine Möglichkeit gibt, historische Daten wieder zu beschaffen.

### 2.1. Missionen und Daten

Das DSDA speichert derzeit Datenprodukte von mehr als dreißig Satellitenmissionen und -sensoren wobei die Ältesten bereits über vierzig Jahre alt sind. Hinzu kommen noch einige flugzeuggestützte Missionen, die besonders experimentellen Charakter haben. Zu den Missionen zählen nationale Missionen, z.B. TerraSAR-X, TanDEM-X, SRTM; europäische Missionen, z.B. ERS, ENVISAT oder METEOSAT sowie internationale Missionen, z.B. TERRA, AQUA oder Landsat.

---

<sup>1</sup>[Dech 2006]

Das gespeicherte Datenvolumen liegt derzeit bei über 2 Petabyte<sup>2</sup>. Hierin enthalten sind sowohl die Rohdaten vom Satelliten (*RAW*- oder *Level-0 (L0)*-Daten) sowie die darauf basierenden, weiter verarbeiteten (*L1-L4*) Daten. Zusätzlich sind in vielen Fällen auch kleinere Vorschaubilder (*Quicklooks*) und bei der Prozessierung der Daten entstandene zusätzliche Dateien gespeichert.

Beim Wachstum des Datenvolumens kann ein Exponentieller Anstieg beobachtet werden. Dies liegt darin begründet, dass nach Möglichkeit alle historischen Produkte archiviert bleiben, zusätzlich jedoch ständig neue Missionen hinzukommen, die aufgrund höherer Auflösungen vielfach auch bedeutend größere Datenmengen produzieren.

### 2.2. Daten- und Informationsmanagementsystem

Das Deutsche Satellitendatenarchiv kann sehr effizient Daten aus verschiedensten Quellen, in unterschiedlichen Größen und Formaten einfügen, archivieren, katalogisieren, deren Weiterprozessierung kontrollieren, Nutzerbestellungen verwalten und bearbeiten sowie anschließend an die Kunden versenden. Diese Datenverarbeitung vom Empfang bis zur Auslieferung wird vom "Daten- und Informationsmanagementsystem" (DIMS) erledigt.<sup>3</sup>

Nachdem eine Satellitenaufnahme getätigt wurde, werden diese Daten über eine *Ingestion*-Komponente an die *Product Library* übergeben. Den Kern von DIMS bilden die *Product Library* und das Archiv. Diese verwalten die Fernerkundungsdaten und deren Metadaten. Dieser Metadatenkatalog wird in Auszügen in die Web-Applikation *EOWEB* gespiegelt. Diese dient als Nutzerschnittstelle zum Auffinden von Produkten, zur Darstellung von Vorschaubildern und zum Tätigen von Bestellungen der eigentlichen Produkte. Die Bestellungen gelangen anschließend ins *Ordering Control*. Hier werden sie nach der Freischaltung durch die Mitarbeiter vom "DFD-Orderdesk" (bei kostenpflichtigen oder sensiblen Daten) automatisch über das *Online-/Offline Product Generation* an den Kunden ausgeliefert. Abbildung 2.1 zeigt eine Übersicht.

Weiterhin sind im DIMS auch die Prozessierungssysteme integriert. Die Schnittstelle zu ihnen stellt die Komponente *Processing System Management* dar und versorgt die Prozessierungssysteme mit Daten und schreibt die neu berechneten Produkte zurück ins Archiv. Über das *Processing Control* lassen sich auch komplexere Verarbeitungsketten steuern.

---

<sup>2</sup>mehr als 2.000.000 Gigabyte

<sup>3</sup>[Dech 2006]

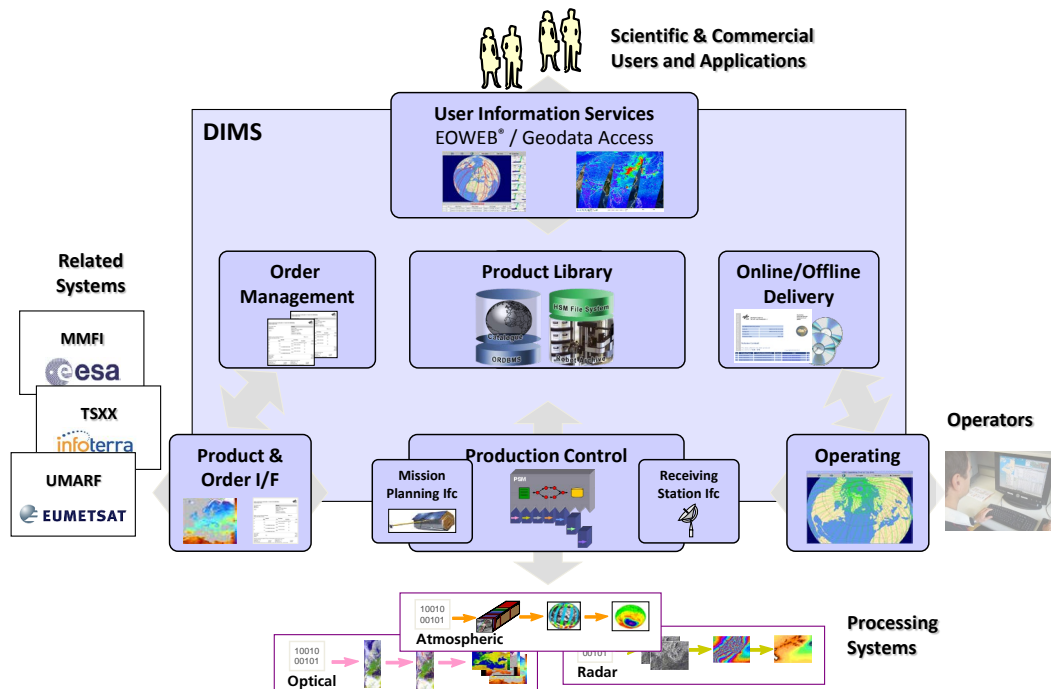


Abbildung 2.1.: DIMS-System-Übersicht

Gesteuert werden alle DIMS-Prozesse mit dem grafischen *Operating Tool*. Hier können für unterschiedliche Anwendergruppen speziell zugeschnittene Ansichten konfiguriert werden, damit die Anwender genau die Systeme sehen und konfigurieren können, die in ihren Zuständigkeitsbereich fallen.

Die DIMS-Software ist eine komplexe, Client-Server-basierte Architektur, die vor ca. 15 Jahren in der Abteilung Informationstechnik des Deutschen Fernerkundungsdatenzentrums des EOC zusammen mit der Lüneburger *Werum AG* entwickelt wurde. Dieses System hat sich in den letzten Jahren zunehmend bewährt, so dass es mittlerweile bei kommerziellen Partnern (Infoterra/Astrium), der Europäischen Weltraumagentur ESA sowie der Südafrikanischen Weltraumbehörde SANSA eingesetzt wird.

### 2.2.1. Product Library

Die Kernkomponenten vom DIMS sind die *Product Library* (PL) und das Archiv und stellen somit das eigentliche, technische DSDA dar. Die PL verfügt über eine große Datenbank mit deren Hilfe alle Metadaten zu den archivierten Produkten abrufbar sind. Somit stellt die Product Library die Schnittstelle zu den meisten anderen

Diensten des DIMS dar. Alle Datentransfers vom und zum Archiv sowie sämtliche Metadatenabfragen zum Katalog erfolgen über die PL. Die PL selbst verfügt über ca. 40 Unterprozesse welche die einzelnen Anfragen verarbeiten und weiterleiten.

### 2.2.2. Archiv

#### 2.2.2.1. HSM

Als Archiv wird ein so genanntes Hierarchisches Speichermanagementsystem (HSM) bezeichnet, welches als komplexes Dateisystem die Datenspeicherung für die Product Library übernimmt. Dieses System besteht aus verschiedenen "Vordergrund-" und "Hintergrundspeichern". Als Vordergrundspeicher dienen mehrere RAID-Systeme von mehr als 70 Terabyte Kapazität, welche als Cache funktionieren. In diese werden über die PL alle Dateien geschrieben und wieder abgerufen. Als Hintergrundspeicher werden zur Zeit drei Bandroboter zum Erstellen mehrerer Sicherheitskopien verwendet. Das eingesetzte *SAM-FS*<sup>4</sup> trennt dabei ebenfalls die Dateiattribute von den eigentlichen Daten. Die Attribute verbleiben dabei dauerhaft im Dateisystem so dass die von UNIX-Systemen bekannten Dateisystembefehle auf alle Daten angewendet werden können.

Wurden nun neue Daten ins Archiv geschrieben, so werden diese automatisch auf zwei Magnetbandkassetten in zwei der Bandroboter geschrieben. Dies geschieht nach frühestens zehn Minuten, um zuerst alle Änderungen an den Dateien abzuwarten, und nach weiteren festgelegten Kriterien (1000 neue Dateien oder fünf Gigabyte neue Daten allerdings spätestens nach einigen Stunden.) Es werden stets zwei Roboter in unterschiedlichen Gebäuden verwendet um im Katastrophenfall (z.B. einem Brand) immer eine der Archivkopien in Sicherheit zu haben. Darüber hinaus werden für die zwei Kopien auch unterschiedliche Bandtechnologien verwendet<sup>5</sup> um vor fehlerhaften Produktionschargen geschützt zu sein. Abbildung 2.2 zeigt das Prinzip von HSM.

Wurden beide Bandkopien erstellt gelten die Daten vorerst als sicher archiviert. Werden nun weitere Daten in den Vordergrundcache geschrieben, so dass der Füllstand die *High Watermark* (z.B. 90%) übersteigt, werden die ältesten, bereits archivierten Dateien *released* (aus dem cache verdrängt). Hierdurch wird der Speicherplatz dieser Dateien frei gegeben, während die Dateiattribute erhalten bleiben.

Fordert daraufhin die PL Dateien an, die bereits released wurden, so sorgt das HSM automatisch dafür, dass diese wieder *gestaged* (vom Band gelesen) werden. Hierzu werden die Roboter angesteuert, um die entsprechenden Bandkassetten aus

<sup>4</sup><http://www.oracle.com/us/products/servers-storage/storage/storage-software/qfs-software/overview/index.html>

<sup>5</sup>zur Zeit "T10000C" mit 5 TB Kapazität und "LTO-5" mit 1,5 TB Kapazität pro Bandkassette



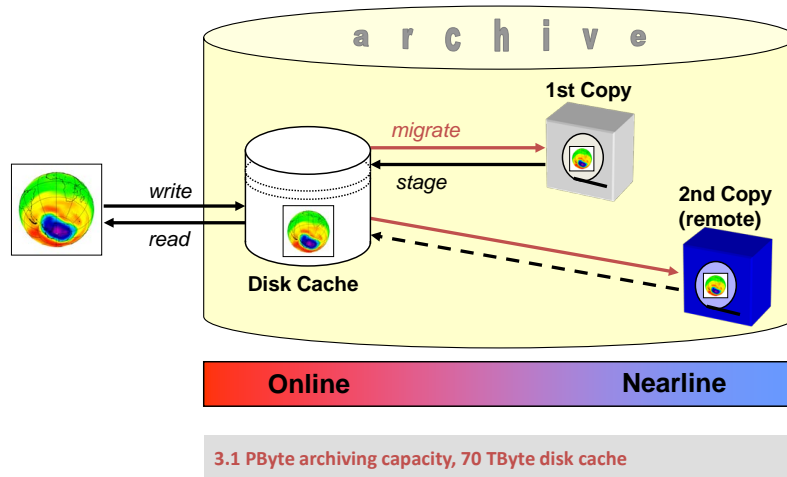


Abbildung 2.2.: HSM-Übersicht

den Lagerstellplätzen in die Bandlaufwerke zu laden. Sie müssen zuerst zum Anfang der Datei auf dem Band spulen, bevor diese gelesen werden kann. Steht die Datei anschließend wieder im Cache zur Verfügung, wird der Product Library dieses signalisiert, woraufhin diese mit der Übertragung dieser Datei über das Netzwerk zum anfordernden System fort fährt.

#### 2.2.2.2. Bänder versus Festplatten

Der Grund, warum bei großen Datenmengen heutzutage immer noch die sehr alte Technik der Speicherung der Daten auf Magnetbänder (Markteinführung ab 1930) erfolgt, sind die Kosten. Festplatten, welche seit 1956 vertrieben werden, haben den großen Vorteil ihrer Geschwindigkeit. Besonders die Zeit, welche verstreicht bis die ersten Daten gelesen werden können (der Latenzzeit) ist sehr gering. Dem gegenüber stehen die hohen Kosten besonders von Laufwerken in "Rechenzentrumsqualität". Hinzu kommt, dass in RAID-Konfigurationen<sup>6</sup> auch nur ein kleinerer Teil der Bruttokapazität der einzelnen Festplatten zur Verfügung steht. Ein weiterer Nachteil ist der Umstand, dass Festplatten ununterbrochen in Betrieb sind, wodurch sich

<sup>6</sup>Zusammenschluss vieler Festplatten um größere Kapazitäten und höhere Datendurchsätze bei gleichzeitig erhöhter Ausfallsicherheit zu erzielen

wiederum die Ausfallwahrscheinlichkeit erhöht (*Mean Time to Failure*.) Während des Betriebs verbraucht jede Festplatte elektrische Energie und produziert eine annähernd gleiche Menge an thermischer Energie, die wiederum mit Hilfe von Klimaanlage abgeführt werden muss.

Magnetbandkassetten auf der anderen Seite verbrauchen nur beim Beschreiben und Lesen im jeweiligen Laufwerk Energie und werden anschließend in einem nur leicht klimatisierten Bereich kostengünstig gelagert. Um den Zugriff auf die Bänder zu automatisieren, werden Robotersysteme zum Laden und Entladen der Laufwerke eingesetzt (Abbildung 2.3). Hierdurch ergeben sich hohe Latenzzeiten, die durch den Ladevorgang durch den Roboter und die Spulzeiten im Laufwerk hervorgerufen werden. Dies ist jedoch im vorliegenden Anwendungsfall in der Regel zu vernachlässigen, da es sich nicht um einen interaktiven Zugriff handelt, sondern um automatisierte Verarbeitungsketten.

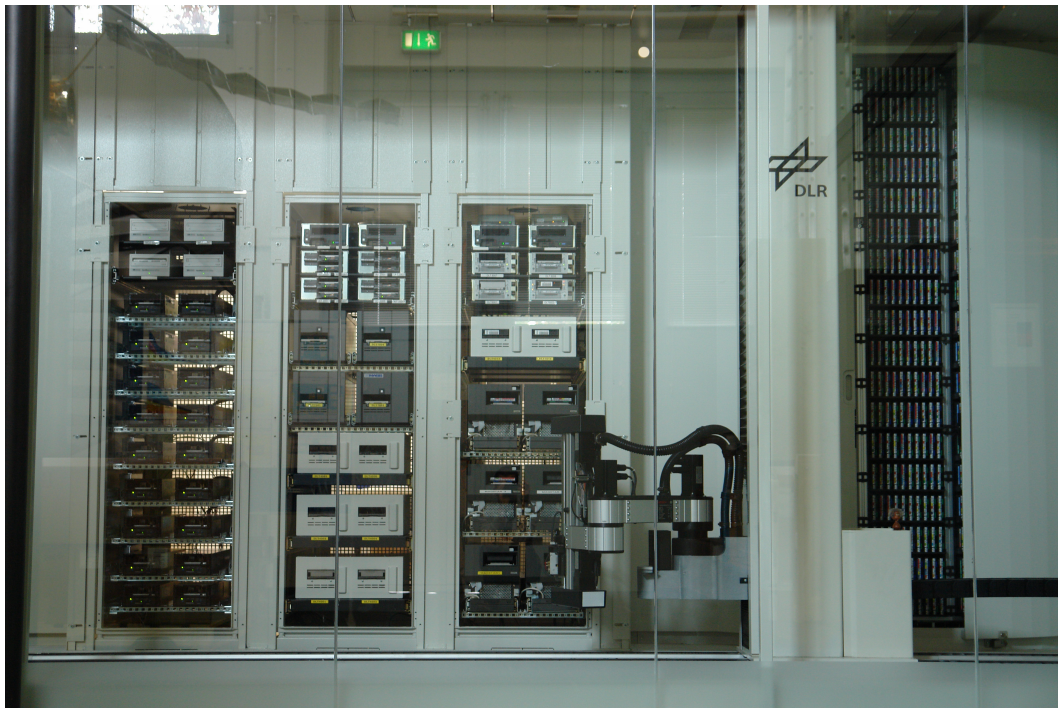


Abbildung 2.3.: Robot Library beim EOC. Links sind die Laufwerke zu sehen, rechts ein Ausschnitt auf die Bandstellplätze, im Vordergrund der Roboterarm.

Beim DSDA werden demnach die Vorteile sowohl von Magnetbändern als auch von Festplatten genutzt, um deren jeweiligen Nachteile zu auszugleichen: Festplatten werden als schneller Vordergrundspeicher (Cache) verwendet. Dieser Cache kann jedoch nur einen kleinen Teil der Daten aufnehmen, wodurch die Kosten hierfür vertretbar sind. Als kostengünstiger Massenspeicher im Hintergrund dienen die Mag-

netbänder, die durch ihre seltenere Nutzung eine vergleichsweise hohe Lebensdauer aufweisen.

### 2.2.2.3. Archivwartung

Auch wenn es sich bei Magnetbandkassetten um langlebige Medien handelt (Hersteller geben oft eine Lebensdauer von 30 Jahren und mehr an), so darf sich besonders im Bereich der Langzeitarchivierung nicht darauf verlassen werden. Die Medien, die Roboter und erst recht die Laufwerke produzieren immer wieder Fehler, die besonders im Laufe der Zeit stark zunehmen. Da im Verhältnis zur Anzahl der Bänder nur relativ wenige Laufwerke zur Verfügung stehen (es sind nur so viele Laufwerke vorhanden, wie zum parallelen Abarbeiten der Anfragen zu einem bestimmten Zeitpunkt benötigt werden)<sup>7</sup>, sind diese nahezu ununterbrochen im Einsatz. Dies führt zu einem hohen Maß an mechanischem Verschleiß bei den heute nötigen hochpräzisen Leseköpfen. So lange die jeweilige Laufwerksgeneration vom Hersteller unterstützt wird, werden diese bei einem Defekt innerhalb weniger Tage vom Hersteller getauscht. Dies ist durch Serviceverträge geregelt. Bei den Robotern verhält es sich ähnlich, wobei diese nicht komplett getauscht sondern vor Ort repariert werden. Zu einem Problem kommt es, wenn der Wartungszeitraum des Herstellers abläuft. Spätestens ab diesem Zeitraum müssen alle Laufwerke und Bandkassetten durch eine neuere Generation ersetzt werden, da auch die Bänder im Normalfall nicht mehr von neueren Laufwerken gelesen und beschrieben werden können.

Wenn es zu Lesefehlern bei einzelnen Bändern kommt, müssen diese rearchiviert werden. Hierbei müssen alle Daten vom jeweiligen Band komplett in den Cache gelesen werden, um anschließend auf eine neues Band zurück geschrieben zu werden. Diese Fehler können sowohl während des normalen Betriebes auftreten, als auch bei routinemäßigen Lesetests.

Wird von einer alten Laufwerks- und Bandgeneration auf eine neue gewechselt, so muss ebenfalls der gesamte Datenbestand migriert werden. Dies ist wichtig, um nicht auf veraltete Bandlaufwerke angewiesen zu sein und während der Migration die Datenintegrität sicherzustellen. Ein positiver Nebeneffekt ist, dass durch die höhere Kapazität von neuen Bandgenerationen auch die Gesamtkapazität des Archivs steigt. Im Deutschen Satellitendatenarchiv wird so eine komplette Migration ungefähr alle sechs Jahre durchgeführt.

---

<sup>7</sup> zur Zeit 14000 Bänder und 39 Laufwerke

### 2.3. Standorte Oberpfaffenhofen und Neustrelitz

Der Großteil des DIMS-Systems befindet sich am Standort Oberpfaffenhofen bei München. Hier sind das Langzeitarchiv, die Web-Schnittstelle EOWEB, das Order-Handling, die Datenauslieferung und weitere Komponenten angesiedelt.

Am Standort Neustrelitz, an der Mecklenburgischen Seenplatte, befindet sich die nationale Empfangsstation. Hier werden nahezu alle über Deutschland gesendeten Daten empfangen. Um diese Daten gleich von Anfang an DIMS-Konform verarbeiten zu können, befinden sich hier ebenfalls eine Product Library und zwei Roboterarchive. Somit können bereits in Neustrelitz viele Prozessierungssysteme die Daten direkt vor Ort weiter verarbeiten. Anschließend werden einige der Daten dann nach Oberpfaffenhofen ins Langzeitarchiv übertragen.

### 2.4. Systemgrenzen

In diesem Abschnitt sollen die Grenzen der einzelnen Subsysteme aufgezeigt werden, das Umfeld beschrieben sowie das in dieser Arbeit betrachtete System mit seinen Grenzen definiert werden.

#### 2.4.1. Beschreibung des Umfeldes

Das DSDA wird von der Abteilung Informationstechnik des Deutschen Fernerkundungsdatenzentrums (DFD) betrieben. Dieses wiederum ist ein Institut des Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR) und bildet zusammen mit dem Institut für Methodik der Fernerkundung (IMF) das *Earth Observation Center* (EOC).

Innerhalb der Abteilung sind die folgenden Teams maßgeblich am Betrieb und der Weiterentwicklung beteiligt: DSDA-Services betreut die operationelle Software-Umgebung, betreibt und konfiguriert die einzelnen Services und hält den Kontakt mit den Nutzern aufrecht. Die Mitarbeiter stellen den *First-Level-Support*. IT-Engineering auf der anderen Seite ist vor allem für die den Diensten zugrunde liegende Hardware zuständig und dient als *Second-Level-Support* in diesem Bereich. Das Team Software-Entwicklung DIMS koordiniert die Entwicklung der Software und erstellt auch viele der Dienste selbst. Die Mitarbeiter leisten daher auch den Second-Level-Support für die DIMS-Services falls das Betriebsteam mit nicht-trivialen Fehlerfällen konfrontiert wird. Als letztes wacht das Qualitätsmanagement über alle Teams und koordiniert die Aktivitäten rund um die ISO-9000-Zertifizierung,

die Form der Dokumentation und die Software-Abnahmen und somit die Übergabe der Dienste in den Betrieb.

Weitere wichtige Beziehungen bestehen zu den Datenlieferanten (Die Abteilungen Nationales und Internationales Bodensegment, europäische und internationale Welt-raumorganisationen (wie der ESA, EUMETSAT, der NASA u.a.) sowie kommerzielle Partner (wie Infoterra/Astrium oder Euromap).

Als Kunden gelten sowohl DLR-interne wie externe Wissenschaftler, Partner aus der Wirtschaft, als auch nationale Behörden, Ämter und Einrichtungen.

### 2.4.2. Blackbox- und Greybox-Betrachtung

Wenn das gesamte DIMS-System als Blackbox betrachtet wird, so sind Datenströme in das System, wie auch wieder hinaus zu beobachten: eingehend sind die Rohdaten von den Satelliten-Empfangsstationen, sowie weiter verarbeitete Daten von den Prozessierungssystemen. Andererseits gehen auch die Kundenbestellungen in das System hinein. Heraus kommende Datenströme sind die Produktauslieferungen an die Kunden sowie die angefragten Rohdaten für die Prozessierungssysteme.

Wird diese Blackbox etwas stärker beleuchtet, so können auch die einzelnen Subsysteme wahrgenommen werden, wie sie bereits unter 2.2 genauer beschrieben wurden. Wird das DSDA im engeren Sinne nur bestehend aus der Product Library und dem Archiv betrachtet, so sind die gleichen Datenströme wie bei der Blackbox-Betrachtung ins System und wieder hinaus zu beobachten: neue Daten werden eingelesen um diese zu archivieren und ältere der Außenwelt zur Verfügung gestellt. Außerdem gibt es jedoch auch eine Vielzahl von Abfragen des *Inventories*, also der Datenbank der Product Library. Diese können interaktive Nutzerabfragen auf den Datenbestand oder Subskriptionen auf neu eintreffende Produkte von Prozessierungssystemen sein.

Darüber hinaus gibt es immer Schnittstellen für die Betreiber des Systems. Dies ist einerseits das Operating Tool und andererseits auch eine als *Monitoring and Alarm* bezeichnete Komponente, welche periodisch die einzelnen Zustände des Systems abfragt und die Operator über Fehler informieren kann.

### 2.4.3. Systemabgrenzung

Je nach gewähltem Betrachtungswinkel ergibt sich ein anderes Umfeld des Systems und somit auch ein anderer Umfang. Der Kunde sieht in der Regel nur das Gesamtsystem ohne tiefen Einblick in selbiges, der DIMS-Operator alle Komponenten auf der Software-Ebene, der IT-Ingenieur die Hardware und Netzwerke und der

## *2. Das Deutsche Satellitendatenarchiv*

---

Software-Entwickler in der Regel nur eine der Softwarekomponenten, diese dafür sehr detailliert.

Diese Arbeit wird im Folgenden das Satellitendatenarchiv als System bestehend aus den beiden Komponenten Product Library und Archiv betrachten.

## 3. Long Term Data Preservation

### 3.1. Einführung

Die Europäische Raumfahrtagentur ESA sieht die Erhaltung von Erdbeobachtungsdaten als Reichtum der Menschheit an und damit als eine Verpflichtung der einzelnen Raumfahrtagenturen. Laut ESA gibt es in Europa und Kanada nur eine schlechte Kooperation auf diesem Gebiet, auch wenn es für die Nutzercommunities sehr von Vorteil wäre. Daher hat sich die Long Term Data Preservation Working Group der ESA das Ziel gesetzt, die Anstrengungen der einzelnen Mitglieder in diesem Bereich zu unterstützen.<sup>8</sup>

Als wichtiges Instrument dienen die *LTDP Common Guidelines*, da sie einen gemeinsamen und umfassenden Ansatz beschreiben, welche Punkte bei der Langzeitarchivierung nötig sind. Anhand derer können sich die einzelnen Mitgliedsorganisationen selbst bewerten. Dies wurde in diesem Kapitel für das EOC, basierend auf der LTDP-Common-Guideline-Version 1.1, durchgeführt.

### 3.2. Bearbeitung

Die folgenden Abschnitte entsprechen genau der Abfolge der LTDP Guidelines. Der einfachen Zuordnung wegen sind die Überschriften direkt in englischer Sprache übernommen worden. Es folgt das Niveau der Richtlinie (Stufe A: Richtlinien mit hoher Priorität; Stufe B: Richtlinien mit mittlerer Priorität; Stufe C: Richtlinien mit niedriger Priorität) sowie eine Beschreibung. Im Anschluss werden die Erkenntnisse in Bezug zum DSDA erläutert. Diese wiederum wurden durch Interviews mit Experten aus den Bereichen Management (Dipl.-Inf. E. Mikusch), EOC-Librarian (M.A. K. Molch), Systemetrieb (Dipl.-Met. P. Seifert, K.-D. Missling), Software Engineering (Dipl.-Inf. S. Kiemle), IT-Engineering (Dipl.-Ing. W. Wildegger), IT-Security (Dipl.-Ing. G. Gräter) sowie Archive-Engineering (der Autor selbst) gewonnen. Am Ende dieses Kapitels befindet sich eine Zusammenfassung des Erfüllungsgrades der Guidelines. Daraus leiten sich Empfehlungen für die zukünftige Umsetzung sowie für die im Anschluss in dieser Arbeit behandelten Themen ab.

---

<sup>8</sup><http://earth.esa.int/gscb/ltdp>



### 3.3. Auswertung

#### 3.3.1. Preserved Data Set Composition

Der erste Teil der LTDP Richtlinien beschäftigt sich mit den zu archivierenden Dateien. Diese sind folgendermaßen klassifiziert:

**1 Preserved Data Set Composition** (Stufe A): Die folgenden Datensätze sind bei den jeweiligen Erdbeobachtungsmissionen aufzubewahren:

**Primary Data:** Primärdaten, oder auch so genannte Roh- oder Level-0-Daten, wie sie vom Satelliten empfangen wurden; sowie höher prozessierte Daten, wie zum Beispiel globale Datenprodukte

Diese werden im Archiv gespeichert und stellen die überwiegende Menge der Daten dar.

**Secondary Data:** Zusatzdaten, wie Satellitentelemetrie, Orbitaldaten, Kalibrierungs- und Validierungsdatensätze sowie die Missionsdokumentation  
Sekundärdaten, welche zur Prozessierung benötigt werden, sind archiviert. Diese wurden entweder direkt vom Satelliten übertragen, beim "Ingestion-prozess" erzeugt oder bei der weiteren Prozessierung. Für die Prozessierung nicht benötigte Satellitentelemetrie ist ausschließlich bei den Satellitenbetreibern (z.B. das GSOC) aufbewahrt.

Die Dokumentation auf der anderen Seite befindet sich nicht im Langzeitarchiv. (siehe Kapitel 5)

**Metadata:** Zusatzdaten zu den Primärdaten. Hierzu zählen unter anderen Ort und Zeit der Aufnahme, Satellitenbezeichnung, Instrumentenname und -konfiguration

Die Metadaten sind in der Datenbank der *Product Library* gespeichert und damit direkt mit den archivierten Satellitendaten verknüpft.

**Browse Images When Generated:** verkleinerte Vorschaubilder (so genannte Quicklooks und Thumbnails)

Diese verkleinerten Versionen der Primärdaten werden, sofern sie generiert werden, parallel zu den Primärdaten im Archiv abgelegt. Die Vorschaubilder werden insbesondere in der Nutzerschnittstelle<sup>9</sup> benötigt und sind daher keine Anwendungsspezifischen Dateiformate wie *HDF*, *netCDF* oder *SAFE*, sondern die weit verbreiteten *TIFF*, *PNG* und *JPEG*.

---

<sup>9</sup>EOWEB



#### **Representation Information, Packaging Information, Preservation Descriptive Information:**

Struktur und Semantik der gespeicherten Daten; beschreibende Informationen zu den gepackten Daten; *Reference, Provenance, Context, Lineage*

Die Struktur und der Kontext der Daten entsprechen der Strukturierung innerhalb der *Product-Library*-Datenbank. Informationen zum Inhalt und Struktur der gepackten Datenpakete werden in den meisten Fällen bei der Auslieferung zum Kunden als *Readme*-Datei mit gesendet. *Reference*-, *Provenance*- und *Lineage*-Angaben zu den jeweiligen Daten werden bisher nicht systematisch erhoben und aufbereitet, sind jedoch bei den anstehenden Missionen stärker berücksichtigt.

Für das Einbringen dieser zusätzlichen Daten und Informationen bedarf es des politischen Willens, da es sich beispielsweise um Fach-, Instituts- oder sogar Länderübergreifende Interessen handelt.

Die Erstellung der Repräsentations- und Abstammungsbeschreibung für alte und aktuelle Missionen ist sehr aufwändig und macht daher wohl nur in einzelnen Fällen Sinn, wenn diese Angaben ohnehin gerade überprüft oder bearbeitet werden. Sehr wichtig und vom Aufwand betrachtet viel einfacher ist die Pflege dieser Daten bei zukünftigen Missionen, wenn bei der Planung Rücksicht darauf genommen wird.

#### **2 Archived Data Format** (Stufe B): Benutzung eines standardkonformen Archivformats für die *Archived Information Packages*

Es wird der EOC-weite Quasi-Standard *Item Information File* (IIF) zur einheitlichen Beschreibung aller Daten im DIMS-Archiv verwendet.

#### **3 Archived Data Exchange Format** (Stufe B): Benutzung eines standardkonformen Datenformats für den Austausch zwischen Archiven

Hier wird ebenfalls der IIF-„Standard“ zum automatischen Austausch der Daten zwischen den Archiven innerhalb des EOC (Oberpfaffenhofen und Neustrelitz), sowie vereinzelt externen kommerziellen und öffentlichen Partnern verwendet.

Positiv am IIF ist, dass bereits darauf geachtet wurde, ein einheitliches Format durchgehend zu implementieren und zu benutzen. Außerdem ist dieses durch die gegebene Verbreitung der DIMS-Software bei einigen anderen Firmen und Organisationen keine kleine Insellösung mehr. Auf der anderen Seite ist ein wenigstens europaweites, wenn nicht sogar weltweites Harmonisieren wünschenswert. Hierfür wurde von der ESA das SAFE-Format spezifiziert. Eine komplette Umstellung des Deutschen Satellitendatenarchivs auf dieses neue

Format ist aber wiederum sehr aufwändig weil jede beteiligte Software geändert und der gesamte Datenbestand umformatiert werden müsste. Hier soll in Zukunft zweigleisig gefahren werden: zum einen wird in Betracht gezogen, für neue Missionen das SAFE-Format zu verwenden, wenn die wissenschaftliche Gemeinde ebenfalls daran interessiert ist. Auf der anderen Seite bietet es sich an, SAFE nur als neue Schnittstelle zu definieren und einzusetzen. Hierzu sind nur wenige Softwarekomponenten anzupassen und die Kompatibilität zu externen Partnern wäre gewährleistet, dem Hauptziel dieses europäischen Standards.

#### 3.3.2. Archive Operation and Organization

##### 1 Reference Model For Archive Operation (Stufe A): Verwendung eines standardisierten Referenzmodells für das Archiv

Der Archivbetrieb basiert auf den Standards des *Open Archival Information System* (OAIS).

##### 2 Operations Procedures (Stufe B): Der Archivbetrieb soll anhand von geprüften und gefestigten Betriebsprozeduren ausgeführt werden

Es sind bereits ein Großteil der Prozeduren für den DIMS-System-Betrieb schriftlich in einem Wiki abgelegt. Diese Dokumentation war auch Grundlage für die ISO-9001-Zertifizierung der *Entity* "DIMS-Operations". Die Art und die Genauigkeit der Dokumentation wurde im Laufe des Zertifizierungsprozesses positiv bewertet. Für den Archivbetrieb sind dort ebenfalls die wichtigsten Prozeduren beschrieben. Diese sollen jedoch noch erweitert werden um z.B. Fehlerfälle und mögliche Lösungsansätze wiederzugeben.

Betriebsprozeduren für die darunter liegende Hardware und Infrastrukturkomponenten, wie zum Beispiel die Netzwerke, das Speichernetzwerk (SAN) und ähnliches ist bislang nicht ausreichend niedergeschrieben beziehungsweise nicht für alle Mitarbeiter einfach zugänglich. Hier ist zu erwarten, dass, im Zuge der ISO-9001-Zertifizierung der *Entity* "IT-Engineering", diese Lücke geschlossen wird.

##### 3 Archive Equipment Maintenance (Stufe A)

Einfache Arbeiten an der Hardware werden im Haus ausgeführt. Darüber hinaus bestehen für alle Hardwarekomponenten Wartungsverträge mit den Herstellerfirmen beziehungsweise spezialisierten Drittfirmen.

**4 Archive Automation** (Stufe B): es ist eine möglichst hohe Archivautomatisierung anzustreben um Operatoreingriffe zu minimieren.

Zum einen arbeitet die DIMS-Software im Normalbetrieb ohne Zutun eines Operators. Auf der anderen Seite werden auf der Archivseite ebenfalls selbsttätige Systeme eingesetzt. Dies ist zum einen das Hierarchische Speichermanagement (HSM). Hierbei werden die Daten einerseits automatisch über schnelle Festplattensysteme auf langsame Massenspeicher (Magnetbänder) übertragen. Andererseits können diese dann genauso transparent wieder ins Dateisystem zurück gespielt werden. Um die Automatisierung konsequent weiter zu führen sind mehrere *Tape-Libraries* im Einsatz, um die Magnetbandkassetten in die jeweiligen Lese- und Schreibgeräte zu laden und wieder zu entladen.

**5 Archive Organisation** (Stufe A): Ziel ist es, eine angemessene Organisationsstruktur basierend auf einer ausreichenden Anzahl qualifizierter Mitarbeiter mit klar verteilten Rollen und Verantwortlichkeiten zu etablieren.

Seit Inbetriebnahme des Archivs sind dedizierte Mitarbeiter für die Installation, Konfiguration, den Betrieb, sowie für die Fehlerdiagnose und deren Behebung zuständig. Als Reaktion auf die starke Zunahme der Nutzung des Archivs durch die aktuellen Satellitenmissionen TerraSAR-X und TANDEM-X wurde außerdem die Anzahl der in diesem Bereich geschulten Mitarbeiter weiter erhöht. Dadurch konnte ein nahezu reibungsloser Betrieb auch während der Abwesenheit einzelner Mitarbeiter sichergestellt werden. Weiteres Verbesserungspotenzial besteht bei der Organisation in Hinsicht auf die Rollenverteilung und Verantwortlichkeiten, sowie bei der Dokumentation von Fehlerfällen<sup>10</sup>.

**6 Archive Legal and Contractual Aspects** (Stufe B): rechtliche Aspekte und vertragliche Regeln sind für das Einlesen, Archivieren und Ausliefern von Datenprodukten festzulegen.

Bei allen neuen Missionen (beziehungsweise bei neuen Produkten) werden über die Rechtsabteilung mit den beteiligten Parteien die rechtlichen Kriterien und die Lizenzen für den Datenzugriff definiert. Diese spiegeln sich zum einen in den Zugriffsregeln für unterschiedliche Benutzergruppen wider. Darüber hinaus werden die Lizenzbestimmungen auch bei allen Auslieferungen an Kunden dem Datenpaket beigelegt. Ein Großteil des im DLR verfügbaren Bildmaterials wird außerdem unter der *Creative-Commons*-Lizenz veröffentlicht, da dieses "nicht der satzungsrechtlichen Verpflichtung des DLR zur Veröffentlichung seiner Forschungsergebnisse"<sup>11</sup> widerspricht.

---

<sup>10</sup>Vgl. Betriebsprozeduren

<sup>11</sup>[http://www.dlr.de/blogs/desktopdefault.aspx/tabid-5921/9755\\_read-503/](http://www.dlr.de/blogs/desktopdefault.aspx/tabid-5921/9755_read-503/)

- 7 Archive Documentation** (Stufe A): das gesamte Archiv und seine Komponenten sind zu dokumentieren und unter Konfigurationsmanagement zu stellen.

Das Archiv ist gut dokumentiert und Änderungen werden per Konfigurationsmanagement erfasst. Hierzu zählen der allgemeine Aufbau, Funktionsweise und Fehlerfälle. Die Dokumentation bedarf allerdings einer Überholung um sie den aktuellen Gegebenheiten anzupassen.

#### 3.3.3. Archive Security

- 1 Archive Security Requirements** (Stufe B): Einhaltung grundlegender internationaler Standards und Richtlinien.

Es werden grundsätzliche Sicherheitsstandards eingehalten. Dies haben Audits auf Basis des DLR-weiten IT-Sicherheitsstandards sowie des deutschen Satellitendatensicherheitsgesetzes<sup>12</sup> bestätigt. Darüber hinaus wurden jedoch bisher keine Anstrengungen unternommen, internationale Standards zu evaluieren um deren Einhaltung zu verifizieren.

- 2 Archive Security Certification** (Stufe C): die Zertifizierung der Archivsicherheit.

Eine Zertifizierung im Bereich Archiv- oder IT-Sicherheit auf der Grundlage von internationalen Standards besteht nicht.

- 3 Controlled Access to Archive Facilities** (Stufe A): es soll eine Zugangskontrolle eingerichtet werden um das physische Eindringen unauthorisierter Personen zu verhindern. Außerdem ist der Zugriff auf Kernfunktionen nur speziellen, überprüften Personen mit entsprechender Sicherheitsfreigabe gestattet.

Sämtliche Rechnerräume sind zugangskontrolliert, so dass der Zutritt nur wenigen bekannten Mitarbeitern und Technikern gestattet ist. Alle Zugänge werden protokolliert. Darüber hinaus müssen sich auch alle Mitarbeiter und Servicetechniker, die Zugang zu den Systemen benötigen, zuvor nach dem Sicherheitsüberprüfungsgesetz des Bundes<sup>13</sup> begutachten lassen.

- 4 Local Risk Mitigation Infrastructure** (Stufe A): Maßnahmen zur Risikominderung und Schutz des Archivs vor externen Einflüssen (Überflutung, Feuer, allgemeine Havarien)

In den Rechnerräumen sind Rauchmelder installiert, die alle Geräte stromlos schalten können, falls es brennen sollte. Außerdem befinden sich Rauchsaugsysteme in den Computerracks, um diese ebenfalls stromlos zu schalten,

---

<sup>12</sup><http://www.gesetze-im-internet.de/satdsig>

<sup>13</sup>[https://bmwi-sicherheitsforum.de/handbuch/text/?fk\\_menu=59](https://bmwi-sicherheitsforum.de/handbuch/text/?fk_menu=59)

sollte z.B. ein Netzteil anfangen zu brennen. In den neuen Rechnerräumen ist zusätzlich eine Argonlöschanlage installiert. Gegen Stromausfälle schützen redundante Unterbrechungsfreie Stromversorgungen. Mit Überflutungen ist im allgemeinen nicht zu rechnen, aber dagegen ist das Archiv insofern abgesichert, als dass es verschiedene Kopien der Daten in unterschiedlichen Gebäuden gibt.

- 5 Protection From External Intrusion (Stufe A):** Verhinderung des Eindringens von Außen zum Stören des Betriebs oder Löschen von Daten.

Alle Bereiche, auch die Kommunikationsinfrastruktur, ist Zugangsgeschützt, so dass nur autorisierte Personen Zugang haben. Deren Zutritt wird zudem protokolliert.

- 6 Controlled Data Access and Products Dissemination (Stufe B):** der Zugriff auf die Daten und die Auslieferung der Daten muss auf bestimmte Kundenkreise beschränkt sein.

Sofern die Daten dem Satellitendaten-Sicherheitsgesetz unterliegen, oder kostenpflichtig sind, erlangen nur ausgewählte Kunden Zugriff darauf. Dies beinhaltet eine vorherige Registrierung, die schriftliche Beantragung des Zugriffs auf die Daten mit Angabe des Forschungsauftrags, die Genehmigung des Antrags durch das Bundesamt für Ausfuhrkontrolle sowie die Beschränkung des Datendownloads auf diese Benutzergruppen inklusive des Zugriffs auf die Daten ausschließlich über gesicherte Verbindungen.

- 7 Information Loss Risk Mitigation Infrastructure (Stufe A):** Einrichtung von Schutzmaßnahmen zum Schutz der Daten und der Hardware vor absichtlichem oder unbeabsichtigtem Informationsverlust.

Wiederum ist der Zugriff nur wenigen, geschulten Personen gestattet und die Daten und die Hardware sind mehrfach redundant ausgelegt, sodass das Gesamtsystem trotz des Ausfalls einzelner Komponenten weiterhin funktioniert.

#### 3.3.4. Data Ingestion

- 1 Data Ingestion Process (Stufe A):** Einfügen neuer Daten anhand von definierten Standards.

Alle Archivoperationen beruhen auf dem "Referenzmodell für ein offenes Archiv-Informationssystem" (OAIS / ISO 14721)

- 2 Metadata Generation (Stufe A): Erzeugung von Metadaten während des Einfügeprozesses. Die entstehenden Metadaten sollen ebenfalls standardkonform formatiert sein.

Die Metadatenerzeugung ist ebenfalls OAIS-konform und geht teilweise über dessen Anforderungen hinaus. Diese Informationen sind in der Product Library in einem eigenen Datenformat (Item Information File) gespeichert, können jedoch beim Exportieren in beliebige andere Formate gewandelt werden.

- 3 Definition of Preserved Data Subsets (Stufe B): Harmonisierung der Inputdaten des Datenlieferanten mit den Standards des Archivs damit alle eingefügten Daten gleichermaßen verständlich und nachhaltig abgelegt werden.

Dieses Kriterium wird nur zum Teil erfüllt, da zwar alle neuen Datenmodelle nach dem gleichen Vorgehen angelegt werden, dieses sich jedoch mit der Zeit verändert und die alten Datenbestände nicht entsprechend aktualisiert werden.

- 4 Routine Quality Check (Stufe A): routinemäßiger Qualitätscheck aller Daten vor dem Einfügen ins Archiv.

Dies wird nur ungenügend erfüllt: es wird lediglich überprüft, ob alle Schlüssel-Parameter und Primärprodukte in die jeweilige Missions-Kollektion abgegeben wurden. Darüber hinaus ist zu prüfen, ob Checksummen eingeführt werden und die Produkte während der gesamten Verarbeitungskette begleiten, so dass die Unversehrtheit der Daten sichergestellt werden kann. Ebenso könnten Plausibilitätschecks auf zum Beispiel die Dateigrößen eingeführt werden, um eindeutig fehlerhafte Produkte von vornherein ablehnen zu können.

- 5 Quality Check (Screening) (Stufe A): ein Screening auf die im Archiv befindlichen Daten, um deren Unversehrtheit zu gewährleisten.

Dies wird in regelmäßigen Abständen vollzogen, indem periodisch alle Daten eingelesen und erneut geschrieben werden. Hierbei helfen die in den Laufwerken integrierten Algorithmen zum Verifizieren der Daten. Dies könnte jedoch mit Checksummen auf PL-Ebene ergänzt werden.

#### 3.3.5. Archive Maintenance

- 1 Archive Data Refreshment (Stufe A): periodische Datenmigration auf geeignete Datenspeichertechnologien um den Datenzugriff sicherzustellen.

Ungefähr alle sechs Jahre wird der gesamte Datenbestand auf neue Medien migriert. Dabei handelt es sich um in der Industrie weit verbreitete und erprobte Technologien.

- 2** Archived Data Formats Description and Converters (Stufe B): Bereitstellung einer formalen Beschreibung der Daten sowie von Konvertern für alte Formate um die Kompatibilität mit anderen Archiven zu gewährleisten.

Eine formale Beschreibung der Daten per *Item Information File* ist in jedem Fall vorhanden. Konverter für alte Formate werden nur in einigen Fällen bereit gestellt. Im Zuge von Reprozessierungskampagnen werden in der Regel auch neuere Datenformate verwendet.

- 3** Archived Data Repackaging / Reformatting (Stufe C): Neuformatierung bzw. erneutes Verpacken der Produkte in aktuelle Formate um die Kompatibilität zu erhöhen. Dieser Vorgang sollte zusammen mit der Archivmigration erfolgen.

Die Produkte werden nicht regelmäßig neu formatiert oder verpackt, dies geschieht nur bei einer Reprozessierung der gesamten Daten einer Mission. Ebenso wenig wird dies bei der Archivmigration durchgeführt, da diese transparent auf der physikalischen Datenebene erfolgt und die Dateien dabei nicht logisch verarbeitet werden.

- 4** Archived Data Duplication: Anlegen identischer Kopien aller Archivdaten nach den folgenden Sicherheitsstufen:

- a** Stufe A: zwei Kopien am selben Standort, jedoch in unterschiedlichen Gebäuden, um Datenverlust zu vermeiden.

Dies wird erfüllt.

- b** Stufe B: zwei Kopien am selben Standort, jedoch in unterschiedlichen Gebäuden, auf unterschiedlichen Technologien um technologiebedingte Fehler zu vermeiden.

Dieses wird ebenfalls erfüllt.

- c** Stufe B: zwei Kopien an unterschiedlichen Standorten um Datenverlust durch Überflutung, Naturkatastrophen oder technologische Risiken zu vermeiden.

Das wird nur teilweise für einen kleinen Teil des Datenbestandes erfüllt (siehe Kapitel 4).

- d** Stufe C: zwei Kopien an unterschiedlichen Standorten und auf unterschiedlichen Technologien.

Das wird nur teilweise für einen kleinen Teil des Datenbestandes erfüllt (siehe Kapitel 4)

- 5** Archive System Components Migration (Hardware) (Stufe A): Durchführung regelmäßiger Hardwareupgrades etwa alle fünf bis sechs Jahre.

Etwa alle sechs Jahre findet zusammen mit der Archivmigration auch ein Upgrade der gesamten im Archiv verwendeten Hardware (außer den Roboter-Systemen) statt.

- 6** Media Readability and Accessibility Tests (Stufe B): Regelmäßige Lesetests auf einen repräsentativen Teil der archivierten Daten.

Etwa alle drei Jahre werden alle archivierten Daten auf Medien des selben Typs rearchiviert um so deren Lesbarkeit zu verifizieren.

- 7** Obsolete Media Disposal (Stufe B): umweltverträgliche Entsorgung gebrauchter Medien.

Gebrauchte Archivmedien werden vom lokalen Abfallentsorger in verschlossenen Behältnissen einer umweltverträglichen Entsorgung zugeführt.

- 8** Archiving System Common Approach (Stufe C): Anstreben eines gemeinsamen Archivkonzeptes mit anderen Organisationen um die Kompatibilität untereinander zu verbessern.

Hierzu findet ein reger Austausch innerhalb und außerhalb der Long Term Data Preservation Working Group statt. Dies umfasst europäische, kanadische, US-amerikanische und neuerdings auch südafrikanische Institute.

#### 3.3.6. Data Access and Interoperability

- 1** Data Accessibility (Stufe A): Bereitstellen und Aktualisieren von Services zum Zugriff auf die gespeicherten Datenprodukte durch:

- Suchmechanismen zum Auffinden gespeicherter Produkte: vorhanden (EOWEB)
- Durchsuchbarer Metadaten- und Quicklook-Katalog: vorhanden (EOWEB)
- Produktgenerierung: vorhanden bei laufenden Missionen
- Produktauslieferung: durch einen Bestellvorgang bzw. *Direct Access*

- 2** Online Access and Delivery (Stufe B): Bereitstellen von Online-Zugriffsmöglichkeiten auf die gespeicherten Daten.

Wird über das EOWEB und FTP bzw. FTPS-Server zur Verfügung gestellt.



- 3** Data Access Conditions (Stufe A): Transparente und gut sichtbare Bedingungen für den Datenzugriff zur Verfügung stellen.

Dies ist nur bei aktuellen Missionen der Fall, jedoch nicht unbedingt einheitlich bei älteren Daten.

- 4** Data Access Interfaces and Delivery Formats (Stufe B): Standardisierung der Schnittstellen und Lieferformate.

Die Auslieferformate sind zum Teil auf Formate standardisiert, wie sie in der jeweiligen Fernerkundungs-Community verwendet werden. Standardkonforme Schnittstellen zwischen anderen Datenzentren werden momentan in Form von OGC-Konformen Schnittstellen wie dem EOWEB-Geoportal oder *Heterogeneous Mission Accessibility*-Erweiterungen (HMA) für die bestehenden Dienste.

- 5** Common Approach for Data Access Systems (Stufe C): Wählen eines gemeinsamen Ansatzes für den Datenzugriff.

Hier gilt das gleiche wie beim vorherigen Punkt.

- 6** EO Products Harmonization (Stufe B): Harmonisierung der Datenauslieferungspakete und Dateiformate für unterschiedliche Missionen.

Dies wird nicht progressiv verfolgt, vielmehr wird hier der Input der Communities vorausgesetzt, um die Lieferpakete entsprechend anzupassen, was systembedingt jedoch einfach umzusetzen ist.

- 7** EO Products Realignment (Stufe C): Anpassung der Lieferpakete alter Daten auf die Standards der aktuellen Missionen.

Hier gilt das gleiche wie beim vorherigen Punkt.

- 8** Archive Search Capability (Stufe C): Erweiterung der Suchmöglichkeiten im Archiv und Implementierung von Möglichkeiten zur Extraktion von *Key Features*.

Es lassen sich bereits alle Metadatenparameter durchsuchen. Die Möglichkeiten zur Extraktion von *Key Features* muss jedoch erst genauer definiert und anschließend umgesetzt werden (z.B. Extraktion von Zeitreihen, *Subsetting*).

#### 3.3.7. Data Exploitation and Re-Processing

- 1** Processing Chains (Stufe A): Erhaltung der Fähigkeit zur Erzeugung von Datenprodukten durch entsprechende Pflege der Prozessierungsketten.

Dies wird durch die Abteilungen Nationales und Internationales Bodensegment gewährleistet, die die meisten der Prozessierungssysteme betreiben.

- 2** Reprocessing (Stufe A): Bereitstellen der Möglichkeit, den gesamten Datenbestand einer Mission neu zu prozessieren, wenn ein neuer, bewährter Algorithmus zur Verfügung stehen sollte. Dabei soll die Möglichkeit bestehen, die alten Produkte zu erhalten, falls dies notwendig ist.

Die Möglichkeit besteht in jedem Fall und wird z.B. bei Atmosphärischen Produkten auch regelmäßig genutzt. Geklärt werden muss in jedem einzelnen Fall jedoch die Finanzierung für ein Reprocessing und den höheren Platzbedarf im Archiv.

- 3** Processing / Reprocessing Capability for Long Term Data Series (Stufe B): Zurverfügungstellen von (Re-)Prozessierungskapazitäten für Langzeit-Datenreihen.

Siehe voranstehenden Punkt.

- 4** Higher Level Applications (Stufe C): Bereitstellen von Referenzdaten zur Unterstützung der Entwicklung von höherwertigen Anwendungen.

Dies wird nicht systematisch betrieben, jedoch in einigen Fällen, wie zum Beispiel der "SAR-Ausbildungsinitiative".

- 5** Earth Observation Data / Products Quality (Stufe A): Sicherstellen der Qualität der Fernerkundungsdaten während der gesamten Missionsdauer.

Dies geschieht wiederum nicht systematisch (siehe 3.3.4.4) jedoch gibt es bei vielen Missionen eine dedizierte *Quality Assurance* die die vorhandenen Daten prüft, bevor sie für die Kunden frei gegeben werden.

- 6** Facilitation of Data Exploitation (Stufe C): Unterstützung der Anwender durch die Implementierung von:

- Datenanpassung durch *Postprocessing*: teilweise vorhanden
- Ausführen von Nutzeralgorithmen auf die Daten: wird bisher nicht unterstützt
- Kombination von Daten von unterschiedlichen Sensoren: wird unterstützt und von einigen Anwendungsabteilungen im EOC betrieben, jedoch nicht direkt von den Betreibern des Datenarchivs

- 7** Information Extraction (Stufe C): Unterstützung der Informationsgewinnung aus Datenprodukten und anschließende Erhaltung dieser Informationen.

Dies wird innerhalb von DIMS nicht unterstützt, jedoch zum Teil in den Projekten *UKIS* und *GDAS*.

### 3.3.8. Data Appraisal and Purge Prevention

- 1 Data Set Description (Stufe A): Erstellung und Pflege einer genauen Beschreibung aller im Archiv vorhandenen Datensätze (dies umfasst mindestens eine Beschreibung der Satellitenmission, des Sensortyps, der gespeicherten Produkte (siehe 3.3.1.1), Zeitspanne und Datenvolumen, physischer Speicherort, Datenformate, Archivverantwortlicher sowie der Anwendungsfälle).

Die Beschreibung ist vorhanden und wird vom *EOC-Librarian*, Katrin Molch, verwaltet.

- 2 Data Appraisal Procedure (Stufe A): falls das Löschen von Datensätzen in Erwägung gezogen werden sollte, ist eine "*Data Appraisal Procedure*" anzuwenden um den Wert der Daten zu untersuchen.

Bisher wurden noch keine operationellen Datensätze gelöscht. Falls dies jedoch geschehen sollte, wird in jedem Fall die Community sowie andere Datenzentren zum Wert der Daten befragt.

- 3 Data Purging Alert Procedure (Stufe A): falls das Löschen von Datensätzen in Erwägung gezogen werden sollte, ist zuvor ein "*Data Purging Alert*" an andere Erdbeobachtungsdatenhalter zu richten, um ihnen die Möglichkeit zu geben, die Datensätze zu übernehmen. In diesem Zusammenhang sind alle relevanten Informationen mit zu übergeben.

Dies würde auf jeden Fall unternommen werden.

- 4 Purging Alert Response Time (Stufe A): Antwort innerhalb eines Monats nach Eingang eines *Data Purging Alerts* wenn das Interesse besteht, die Verantwortung für die Daten zu übernehmen.

Dieses Vorgehen ist so akzeptiert.

- 5 Archived Data Handover (Stufe A): Unterstützung des Übergabeprozesses an mögliche Interessenten, die auf einen *Purging Alert* geantwortet haben.

Dieses Vorgehen ist so akzeptiert.

## 3.4. Bewertung

Im Anhang A.2 ist die tabellarische LTDP-Guideline zu finden. Hier sind auch Werte für die Wichtigkeit aus Sicht des EOC und den zu erwartenden Aufwand zur Umsetzung der jeweiligen Richtlinie zu finden. Die Erstellung einer Bewertungsmatrix gestaltete sich dennoch schwierig, weil keine Werte für die Wichtigkeit aus Sicht

Tabelle 3.1.: Grad der Erfüllung für die LTDP-Guideline-Kategorien

Category	Number of Guidelines	Compliance	Compliance [%]
A	26	23.2	89
B	17	12.1	71
C	10	4.8	48
	53	40.1	76

Tabelle 3.2.: Grad der Erfüllung der LTDP-Guideline-Themenbereiche

Guideline Theme	Number of Guidelines	Compliance	Compliance [%]
Preserved Data Set Composition	3	0.7	23
Archive Operation and Organization	7	5.9	84
Archive Security	7	5.5	89
Data Ingestion	5	4.0	80
Archive Maintenance	11	8.6	78
Data Access and Interoperability	8	5.6	70
Data Exploitation and Re-Processing	7	4.8	69
Data Appraisal and Purge Prevention	5	5.0	100
	53	40.1	76

der ESA vorhanden sind. Dadurch erschienen die Ergebnisse einer Bewertungsmatrix sehr willkürlich, weswegen auf die Resultate verzichtet werden musste. Anstelle dessen wird ein anderes Vorgehen gewählt: in der Tabelle 3.1 ist der Erfüllungsgrad der einzelnen Kategorien zusammen gefasst. Dieser ergibt eine sehr hohe Stufe der Einhaltung besonders der Kategorien A und B. Insgesamt werden die Richtlinien zu gut drei Vierteln erfüllt. Und obwohl nahezu die Hälfte von Level C erfüllt wird, wird Level A noch nicht vollständig erfüllt. Bei dem Ziel, die Common Guidelines möglichst komplett umzusetzen, sollten zuerst alle Kategorie-A-Richtlinien erfüllt werden. Bei den Leveln B und C sollte zuerst abgewogen werden, ob die jeweiligen Richtlinien für das EOC relevant sind.

Die Tabelle 3.2 listet nun den Erfüllungsgrad für die einzelnen Themenbereiche auf. Hier ist ebenso ein gleich hohes Niveau bei nahezu allen Bereichen zu sehen. Eine Ausnahme stellt besonders der erste Themenbereich dar. Dies liegt zum großen Teil an der Datenpolitik anderer Institute bezüglich Übergabe und zentraler Speicherung sämtlicher zu einer Mission gehörender Daten und Dokumentationen.

#### 3.4.1. Auswahl der zu behandelnden Themen

Da der erste Themenbereich der LTDP-Common-Guidelines im vorherigen Abschnitt besonders schlecht abgeschnitten hat, was den Erfüllungsgrad betrifft, soll in Kapitel 5 näher auf die Möglichkeit der Speicherung von Missionsdokumentation direkt

in der Product Library und damit im zentralen Archiv eingegangen werden. Die Möglichkeit zur Speicherung von Satellitentelemetrie besteht ohnehin schon. Dies wäre lediglich eine politische Entscheidung, wo diese Daten archiviert werden.

Zuvor soll jedoch im Kapitel 4 die Möglichkeit, die Archive in Oberpfaffenhofen und Neustrelitz zusammen zu legen, näher betrachtet werden. Da die Voraussetzungen hierfür vermutlich bereits gegeben sind, wäre dies eine gute Gelegenheit, eine weitere C-Kategorie zu erfüllen. Außerdem ist dieses Themengebiet dem Autor besonders vertraut.

## 4. Zusammenlegung der Archive in Oberpfaffenhofen und Neustrelitz

Dieses Kapitel soll die Möglichkeit untersuchen, die beiden Archive an den Standorten Oberpfaffenhofen und Neustrelitz zusammenzulegen. Dies würde bedeuten, dass alle Daten von den beiden Standorten am jeweils Anderen gespiegelt werden. Hierfür sollen die Rahmenbedingungen erläutert werden, die Vorzüge beschrieben sowie mögliche Umsetzungen diskutiert werden.

### 4.1. Voraussetzungen

Wie im Kapitel 2 beschrieben, gibt es bereits an den beiden Standorten Oberpfaffenhofen und Neustrelitz jeweils eigenständige Archivinfrastrukturen. Diese bestehen aus jeweils einem eigenständigen Archivroboter für jeweils eine der zwei Archivkopien. Weiterhin gibt es an beiden Orten einen Archivserver, auf dem die HSM-Software SAM-FS für die Archivierung sorgt und auf dem außerdem die archivnahen Services der Product Library (*Archive Areas*) untergebracht sind. Diese Archive Areas betreuen in der Regel einen bestimmten Teil des Archivcaches, der für eine spezielle Satellitenmission zuständig ist und sorgt für das Archivieren und Stagen der einzelnen Dateien sowie für die Übertragung per FTP über das Netzwerk.

Bei der Konzeption der beiden Archive wurde bereits darauf geachtet, dass möglichst gleiche Hardware, in jedem Fall jedoch kompatible Hardware verbaut wurde. So handelt es sich bei den Servern um Hardware der Firma Sun/Oracle, auf denen das Solaris-Betriebssystem arbeitet. Dies ist die Voraussetzung zum Betrieb des SAM-FS. Darüberhinaus sind auch die Robotersysteme jeweils vom gleichen Fabrikat (Storagetek/Oracle SL8500 und Quantum i6000). Auch die eingesetzten Bandlaufwerke sind identisch (siehe Kapitel 2.2.2.1).

Neben den Archivservern gibt es außerdem auch dedizierte Server, auf denen die Datenbanken der beiden PLs laufen sowie weitere PL-Services zum Abfragen und Bearbeiten der Datenbankeinträge. Die beiden Product Libraries laufen zwar eigenständig, sind jedoch locker miteinander gekoppelt, so dass gegenseitige Abfragen auf die jeweiligen Datenbestände möglich sind. Darüber hinaus ist bereits der "transparente PL-Zugriff" implementiert. Dabei handelt es sich um die Möglichkeit für

Client-Systeme eine beliebige der beiden PLs zu kontaktieren und in jedem Fall die richtigen Antworten und Daten zu bekommen, auch wenn diese nicht direkt von der angefragten PL verwaltet werden.

Für die Anbindung der beiden Standorte sind das Deutsche Forschungsnetz sowie T-Systems zuständig. "Das Deutsche Forschungsnetz (DFN) ist das von der Wissenschaft selbst organisierte Kommunikationsnetz für Wissenschaft und Forschung in Deutschland. Es verbindet Hochschulen und Forschungseinrichtungen miteinander und ist nahtlos in den europäischen und weltweiten Verbund der Forschungs- und Wissenschaftsnetze integriert. Über mehrere leistungsstarke Austauschpunkte ist das DFN ebenfalls mit dem allgemeinen Internet verbunden."<sup>14</sup> T-Systems ist auf der anderen Seite damit beauftragt, IT-Basis-Services für das DLR zur Verfügung zu stellen. Hierzu zählt neben der Betreuung der Arbeitsplatzrechner und der E-Mail- und Webserver auch die Netzwerkanbindung der DLR-Standorte an das deutschlandweite Forschungsnetzwerk. Während der weitaus größere Standort Oberpfaffenhofen mit derzeit nominell 600 MBit/s ans DFN angebunden ist, verfügt Neustrelitz über einen 300-MBit/s-Anschluss.

## 4.2. Vorteile

Die Vorteile für eine Zusammenlegung der beiden Archive sind vielfältig. Ein wichtiger Punkt ist die Datensicherheit: bei der Spiegelung des gesamten Datenbestandes des Satellitendatenarchivs über beide Standorte wird eine viel größere räumliche Trennung der Archivkopien erzielt, als dies momentan durch die Unterbringung der Archivkopien in unterschiedlichen Gebäuden am jeweils selben Standort möglich ist. Dies würde die Datenverfügbarkeit selbst bei einem Totalausfall eines der Standorte (sei es durch physische Zerstörung oder Nichtverfügbarkeit von Energie- oder Netzwerkversorgung) ermöglichen und zusätzlich auch die Zugriffsgeschwindigkeit auf die Produkte erhöhen, indem Anfragen von räumlich getrennten Quellen auf das nächste zur Verfügung stehende Archiv geleitet werden. Daraus ergibt sich auch eine Lastverteilung, die hierdurch ermöglicht wird.

Ein großer betriebswirtschaftlicher Faktor wäre eine mögliche Kosteneinsparung da pro Standort weniger Hardware benötigt werden würde. Derzeit sind jeweils zwei Robotersysteme notwendig, um die geforderte räumliche Trennung der Archivkopien zu gewährleisten. Bei einer Zusammenlegung wäre nur noch jeweils ein Roboter mit entsprechend hoher Kapazität nötig. Die bereits vorhandenen Robotlibraries sind darüber hinaus auch bereits in der Lage mit wenigstens zwei verschiedenen

---

<sup>14</sup><https://www.dfn.de>

Medientypen umgehen zu können. Somit kann auch diese Stufe der Datensicherheit gewährleistet werden (siehe Kapitel 3.3.5).

Und nicht zuletzt würde diese Initiative auch eine strukturelle Stärkung des Standortes Neustrelitz darstellen. Der mit nur 70 Mitarbeitern um Größenordnungen kleinere DLR-Standort bekäme die Chance, neben der bereits wichtigen Empfangstätigkeit ein weiteres großes Standbein aufzubauen anstatt wie bisher nahezu ausschließlich eine Durchlaufstation für Erdbeobachtungsdaten zu sein.

### 4.3. Mögliche Lösungsansätze

In diesem Abschnitt sollen nun drei mögliche Ansätze präsentiert werden, wie die Archive verbunden werden könnten.

#### 4.3.1. Nutzung des DIMS-Services UIF

Der Dienst *User Information Services Interface Framework* (UIF) ist ein DIMS-Service, welcher dazu benutzt wird, Metadaten und Produkte zwischen der Product Library und anderen Diensten auszutauschen. Hierzu können Regeln definiert werden, ob und wie die Metadatenstruktur mit Hilfe einer *Stylesheet*-Transformation geändert werden soll und welche Dateien vom gesamten Datenprodukt übertragen werden sollen. Hierzu können entweder *Timer* (die Ausführung erfolgt regelmäßig nach bestimmten Zeitabständen) oder *Trigger* (die Ausführung erfolgt sobald die Triggerbedingung erfüllt ist (z.B. neues Produkt vorhanden)) in der Quellen-PL definiert werden.

Das UIF wird bereits an mehreren Stellen verwendet, so zum Beispiel um neue Produkte automatisch aus der PL in den Web-Katalog EOWEB für die Kunden zu veröffentlichen. Und selbst zwischen den beiden Product Libraries in Oberpfaffenhofen und Neustrelitz gibt es bereits eine UIF-Instanz. Diese transferiert derzeit die Rohdaten der TerraSAR-X-Mission nach dem Empfang in Neustrelitz ins Langzeitarchiv nach Oberpfaffenhofen.

Die Grundfunktionalität ist also bereits abgedeckt. Limitierende Faktoren sind die Größe des Archives in Neustrelitz sowie die Netzwerkbandbreite an beiden Standorten. Dies wird im Kapitel 4.4 noch genauer betrachtet.



#### 4.3.2. Verwendung eines Cluster-Dateisystems

Eine andere Variante, die auf einer tieferen Systemebene ansetzt, ist die Verwendung eines Cluster- oder Cloud-Dateisystems (Cluster-FS). Ein Cluster-FS ist ein verteiltes Dateisystem, dass entweder in einer lokalen Umgebung (einem Cluster) oder auch über weite Entfernungen hinweg (die "Cloud") zwei Dateisysteme synchronisieren kann, sie transparent von beiden Seiten zugänglich machen und gleichzeitig die Größe skalieren kann. Somit ist der gemeinsame Zugriff an beiden Standorten gewährleistet. Um die Daten jedoch überall archivieren zu können, müssen sie allerdings zuerst komplett synchronisiert werden. Der Vorteil der Cluster-FS hierbei ist die "Kenntnis" des unzuverlässigen Internets zwischen ihnen. Sie können die Netzwerk-Bandbreite viel effizienter ausnutzen als es das einfache FTP kann und selbst gegen Ausfälle sind sie robust und können die Zeit überbrücken.

Wenn auf diese Weise die Dateien Synchronisiert wurden, müssen jedoch zusätzlich noch die Metadaten in den PLs aktualisiert werden, damit diese auch vom neuen Datenbestand Kenntnis erlangen. Hierfür kann wiederum das UIF verwendet werden, das sich in diesem Fall jedoch nur um die Metainformationen kümmern muss.

#### 4.3.3. Offline-Transport der Daten

Eine dritte Möglichkeit, große Datenmengen über weite Entfernungen zu transportieren ist die Nutzung einer Spedition. Diese Variante wird bereits bei abgelegenen Empfangsstationen in der Antarktis und in Kanada genutzt. Hierbei werden die empfangenen Daten zuerst lokal zwischengespeichert und eine weitere Bandkopie für den Transport angefertigt. Diese wird dann bei Gelegenheit z.B. mit einem Schiff (aus der Antarktis) auf den Weg ins zentrale Archiv verfrachtet. Wenn die Daten dort erfolgreich eingelesen wurden, kann die lokale Kopie wieder gelöscht werden. Dieser Vorgang kann viele Wochen dauern.

Bei der Verwendung dieser Variante innerhalb Deutschlands bieten sich zwei Möglichkeiten an: entweder die Erstellung einer Bandkopie, die nach dem physischen importieren ins Archiv des anderen Standorts direkt verwendet werden kann ohne noch einmal gelesen und erneut geschrieben werden zu müssen. Oder es werden Festplattensysteme mit ausreichender Speicherkapazität verwendet. Diese können schnell beschrieben werden und "am Stück" per Spedition transportiert werden. Allerdings müssen in diesem Fall die Daten anschließend noch auf die Magnetbänder im Zielarchiv migriert werden.

Wenn davon ausgegangen wird, dass ein Speichersystem mit 100 Terabyte Kapazität innerhalb von zwei Wochen beschrieben, transportiert und wiederum ausgelesen

Tabelle 4.1.: Mehrkosten für die Netzwerkanbindung

Standort	Bandbreite	Kosten/Jahr	Kosten/Monat	Mehrkosten/Monat
<b>Oberpfaffenhofen</b>				
aktuell	600 Mb/s	155.700 €	12.975 €	
zukünftig	1000 Mb/s	207.600 €	17.300 €	4.325 €
<b>Neustrelitz</b>				
aktuell	300 Mb/s	103.800 €	8.650 €	
zukünftig	600 Mb/s	155.700 €	12.975 €	4.325 €
<b>Gesamt</b>				<b>8.650 €</b>

wird, was ein realistischer Wert wäre, so ergibt sich daraus rein rechnerisch eine Transferleistung von 700 MBit/s.

Auch wenn dieses Verfahren sehr aufwändig für den Normalbetrieb anmutet, so könnte es eine gute Möglichkeit für die initiale Synchronisierung beider Archive darstellen.

#### 4.4. Netzwerkanbindung und Kosten/Nutzen-Betrachtung

Um die erforderliche Transferleistung zu ermitteln wird vom gegenwärtigen maximalen monatlichen Zuwachs des Archivvolumens in Oberpfaffenhofen, dem größten Archiv, von 100 Terabyte ausgegangen. Dies entspricht einer durchschnittlichen Transferrate von 325 MBit/s. Das liegt bereits oberhalb der gesamten Bandbreite für Neustrelitz und entspricht der Hälfte der Bandbreite für Oberpfaffenhofen (siehe Kapitel 4.1) Um die anderen Institute besonders in Oberpfaffenhofen nicht auszubremsen ist also an beiden Standorten ein Ausbau der Netzwerkanbindung auf 600 MBit/s bzw. 1000 MBit/s nötig. Die Mehrkosten wird in diesem Fall wohl das Institut DFD übernehmen müssen. Tabelle 4.1 listet die Kosten für die unterschiedlichen Netzwerkanbindungen der Standorte auf<sup>15</sup>. Die monatlichen Mehrkosten ergeben demnach insgesamt 8.650 Euro. Es muss jedoch davon ausgegangen werden, dass der Dienstleister T-Systems einen weiteren Aufschlag verlangt, der zum gegenwärtigen Zeitpunkt nicht bekannt ist.

Um die Kosten in Relation setzen zu können, wird der aktuelle Preis für den Neuaufbau eines Roboters, wie er derzeit durchgeführt wird, herangezogen. Dieser beinhaltet den Roboter, die Laufwerke, die SAN-Anbindung und weitere erforderliche Bauteile sowie Umbaumaßnahmen inklusive dem Abschluss eines Servicevertrages für die Hardwarewartung. Dieser Preis beläuft sich auf ungefähr 1,5 Millionen Euro.

<sup>15</sup>Quelle: <https://www.dfn.de/dienstleistungen/dfninternet/entgelte>

Darin enthalten sind die 13% für den jährlichen Wartungsvertrag, der die folgenden Jahre jeweils zusätzlich anfällt.

Falls es bei diesem Preisunterschied bleibt (T-Systems keinen erheblichen Preisaufschlag verlangt) erübrigt sich eine genaue Amortisationsrechnung! Wenn davon ausgegangen wird, dass die neu beschaffte Hardware nicht modernisiert wird und gleichzeitig die Kosten für die Netzwerkanbindung nicht sinken sollten, ergibt sich eine Amortisation erst nach über 15 Jahren. Und hierin sind noch keine Abschreibungskosten enthalten, die ebenfalls bei der Miete des Internetanschlusses weg fallen.

### 4.5. Schlussfolgerung

Leider liegen keine genauen Kosten des Providers T-Systems für die Aufstockung des Internetanschlusses vor. Werden jedoch die Kosten des Dienstinhabers Deutsches Forschungsnetz zugrunde gelegt, kann von einer deutlichen Kosteneinsparung durch die Nutzung des jeweils anderen Standortes zur Erzeugung der zweiten Archivkopie ausgegangen werden.

Weiterhin sind eventuell noch nicht alle Kosten berücksichtigt. Dies betrifft beispielsweise die Erweiterung des Archivs besonders in Neustrelitz. In Oberpfaffenhofen erübrigt sich dies wegen des ohnehin größeren Ausbaus vermutlich.

Die obige Berechnung hat auch nur das derzeitige maximale Datenvolumen zugrunde gelegt. Besonders für die initiale Synchronisation der Datenbestände reicht die ermittelte Bandbreite nicht aus, oder die Synchronisation würde sich über viele Jahre hinziehen, wenn ausschließlich die verbleibende Bandbreite genutzt wird. Erfolg versprechend scheint für den einmaligen Datenaustausch die Nutzung einer Spedition zu sein.

## 5. Speicherung von Missionsdokumenten in der Product Library

Die Langzeitarchivierung kümmert sich in erster Linie um die physische Datenspeicherung. Auf dieser Ebene ist das Datenformat, Nullen und Einsen, stets gleich geblieben. Durch die Speicherung von mehreren Kopien und die Verwendung von Checksummen kann sichergestellt werden, dass die Daten unverändert viele Jahrzehnte überstehen können. Der Wechsel von einer Speichertechnologie auf eine andere gestaltet sich ebenso einfach, da die Bits lediglich komplett gelesen und neu geschrieben werden müssen.

Ein Problem bei der Speicherung großer, heterogener Datenmengen über einen großen Zeitraum hinweg ist allerdings die Interpretierbarkeit dieser Daten nach langer Zeit. Bereits nach wenigen Jahren ändern sich die Werkzeuge zum Bearbeiten der Daten, werden die Dateiformate aktualisiert und die Struktur der Daten neu geordnet. So kann es sehr schwer sein, die Daten erstens zu finden, zweitens zu lesen und drittens zu verstehen.

### 5.1. Emulation, Migration, Description

Hier hat sich das Paradigma "*Emulation, Migration, Description*" (Emulation, Migration, Beschreibung) von David Giaretta durchgesetzt<sup>16</sup>. Dieses besagt zum einen, dass die alten Werkzeuge zum Beispiel mit Hilfe einer Virtuellen Maschine emuliert werden können. Dies macht es aber schwierig, Daten aus verschiedenen Quellen zu kombinieren. Andererseits können die Daten auch regelmäßig in neuere Versionen konvertiert werden um mit modernen Werkzeugen bearbeitet werden zu können. Hierzu müssen jedoch große Datenmengen mit vielen unterschiedlichen Konvertern bearbeitet werden. Die wohl universellste Variante ist die genaue Beschreibung der Daten, welche in jedem Fall angewendet werden sollte.

Selbst wenn die Produkte mit aktuellen Werkzeugen geöffnet werden können, heißt dies noch nicht, dass der Inhalt dieser Daten selbsterklärend ist. Wichtige Kenntnisse zur Interpretation von Fernerkundungsdaten sind zum einen der genaue Zeitpunkt

---

<sup>16</sup>[Giaretta 2011]

und der geographische Ort, an dem sie aufgenommen wurden. Darüber hinaus ist das Wissen über die Spektralbereiche wichtig. Hierzu muss eventuell noch der genaue Aufbau des Sensors auf dem Satelliten bekannt sein.

Aus diesem Grund hat die *Long Term Data Preservation Working Group* der ESA mit dem *Earth Observation Preserved Data Set Content*<sup>17</sup> eine Empfehlung zur Speicherung wichtiger Daten und Dokumente nach Ablauf einer Satellitenmission herausgegeben.

## 5.2. Geeigneter Speicherort

### 5.2.1. Dedizierter Missionsserver

Derzeit wird für die DLR-Missionen TerraSAR-X und TANDEM-X jeweils ein eigener Server für die Dokumentenverwaltung betrieben. Dies hat den großen Vorteil, dass alle Dokumente die jeweilige Mission betreffend von allen Instituten, Abteilungen und Teams an zentraler Stelle gespeichert sind. Auf diese Weise können, den nötigen Überblick vorausgesetzt, die verschiedenen Versionen von Beschreibungen bestimmter Funktionen gezielt abgerufen werden.

Der Nachteil ist, dass so ein spezieller Server von Projektmitteln finanziert wird und somit der Betrieb nach Missionsende ungewiss ist. Wahrscheinlich werden die Dokumente nach dem Ende einer Mission noch weitere Jahre vorgehalten, jedoch ist einige Jahrzehnte nach dem Ende nicht mehr davon auszugehen.

### 5.2.2. Zentrales Dokumentenmanagement

Das zentrale IT-Management des DLR strebt derzeit ein großes Dokumentenmanagementsystem für das gesamte DLR an. Der Vorteil liegt darin, dass eine zentrale Speicherung einzelne Projekte von dieser Aufgabe entbindet und so auch die Datensicherheit besser gewährleistet werden kann. Ein großer Nachteil ist die Trennung der Daten und Dokumentation, so dass die wichtige Verknüpfung der Beiden (siehe Einleitung) schwer wird.

### 5.2.3. Speicherung direkt in der Product Library

Aus diesen Gründen liegt die Speicherung auch der Dokumentation im *Daten- und Informationsmanagementsystem* (DIMS) nahe. Hierzu ist es weniger erforderlich,

---

<sup>17</sup>[LTDP 2012]

die genaue Versionshistorie aller zu einer Mission entstandenen Dokumente bereitzuhalten, als vielmehr die wichtigsten Dokumente, zum Verständnis der jeweiligen Mission und deren Daten im Archiv, in der jeweils neuesten Version.

Dies könnte einerseits durch einen reinen Verweis innerhalb der Metadateninformationen in der PL auf den tatsächlichen Speicherort der Dokumente (z.B. ein zentrales Dokumentenmanagementsystem) erfolgen. Hierdurch entstehen jedoch externe Abhängigkeiten, die allen Beteiligten Personen bekannt sein müssen, damit diese gepflegt werden können, um immer einen aktuellen Stand zu besitzen.

Auf der anderen Seite lassen sich die wichtigsten, zum Verständnis der Mission erforderlichen Dokumente auch direkt in der PL und dem Archiv ablegen. Dies könnte zum Beispiel am Ende einer Mission geschehen, wenn die Unterlagen einen finalen Stand erreicht haben. Auf diese Weise sind die Daten und die zum Verständnis nötigen Beschreibungen stets beieinander im selben Archiv und direkt miteinander verknüpft, um deren Wiederauffindbarkeit zu erhöhen.

### 5.2.3.1. Konfiguration eines neuen Dokumenten-Produkttyps

Anhand der folgenden Bildschirmfotos soll eine einfache, beispielhafte Konfiguration eines neuen Produkttyps in der Product Library für Dokumente beschrieben werden. Mit Hilfe des Operating Tools wurde unter *Configuration/Space Management* ein neuer Bereich "DemoDocuSpace" zum Testen des neuen Produkttyps angelegt (Abbildung 5.1).

Unter Zuhilfenahme des PL *Space Model Designers* wurde als nächstes ein neuer *ComplexItemType* mit der Bezeichnung "ProductTypeDocumentation" angelegt. Dieser leitet sich vom übergeordneten Typ *Product* ab und erhielt beispielhaft die Attribute *documentTitle*, *documentDate* und *revision* (Abbildung 5.2). Der Quelltext des resultierenden *MetaItemFile* ist im Anhang A.1 abgedruckt.

Jetzt steht der neue Produkttyp "ProductTypeDocumentation" bereit (Abbildung 5.3) und kann für unterschiedliche PL-*Collections* (Missionen) verwendet werden.

### 5.2.4. Offene Punkte

Gezeigt wurde, wie eine generische Produktdefinition für Dokumente aussehen könnte. Die Metadaten können beliebig um weitere Attribute wie Autor, Beschreibung oder ähnliches erweitert werden. Hierzu sollten standardisierte Ansätze<sup>18</sup> verwendet werden. Außerdem wäre denkbar, diese Informationen direkt in die Metadaten der

---

<sup>18</sup>ISO 19115, INSPIRE

## 5. Speicherung von Missionsdokumenten in der Product Library

The screenshot displays the 'Product Collection Browser' application. The top navigation bar includes tabs for 'Product Queries', 'Local Products', 'Configuration', and 'Monitoring & Control'. The 'Configuration' tab is active, showing a 'Processing Systems Registry' table with columns for 'Nr.', 'Name', 'Accessible', 'Defined', 'Compiled', 'Online', 'Description', 'Creation Date', 'Updated', 'Author', and '100/100'. Below the table, there are buttons for 'Now' and 'Edit'. The main configuration area is divided into several sections: 'Characteristics' (with fields for name, description, creation date, and expiration date), 'Access' (with checkboxes for accessible, online, model defined, and model mapped), 'Space Hierarchy' (with a tree view showing 'ExtendedProductSpace' as the parent and 'BaseProductSpace' as a child), 'Version' (with fields for version, date, creator, and comment), and 'JDBC URL' (with a text field for the database connection string).

Abbildung 5.1.: Konfiguration des neuen Bereichs "DemoDocuSpace"

jeweiligen Collection zu schreiben. Beide Ansätze sind bereits in der PL vorgesehen gewesen, wurden jedoch aufgrund der Komplexität der ISO-Norm wieder fallen gelassen. Hier wäre es sinnvoll, die genauen Anforderungen erneut zu diskutieren und anschließend nur die wichtigsten Elemente zu implementieren, so dass sich der Pflegeaufwand auf ein Minimum beschränkt (ähnlich dem einfachen Beispiel in diesem Kapitel).







## 6. Fazit und kritische Bewertung

In den voran gegangenen Kapiteln wurde das DSDA und seine Komponenten genauer beschrieben. Anschließend wurde es anhand der LTDP-Common-Guidelines der ESA bewertet, um den Grad der Erfüllung der einzelnen Themenbereiche zu bestimmen. Daraufhin wurden zwei der Themen ausgewählt, um sie näher zu betrachten. Hierbei handelte es sich um die Zusammenlegung der Archive an den beiden Standorten Oberpfaffenhofen und Neustrelitz sowie um eine Möglichkeit, Missionsdokumente neben den Daten ebenfalls in der Product Library abzuspeichern.

Bei der Auswahl der Bewertungsgrundlage für das DSDA wurde auf die LTDP-Common-Guidelines zurück gegriffen, weil sich dies durch die Mitgliedschaft des DLR in der LTDP-Working-Group anbot. Darüber hinaus ist dem Autor keine andere adäquate Studie bekannt, die dieses Thema so umfangreich umreißt und nicht nur tiefgründig auf Einzelaspekte eingeht. Bei der Diskussion der Ergebnisse besteht noch Potenzial, diese geeigneter zusammen zu fassen und anhand einer Bewertungsmatrix eine Prioritätenliste der Bearbeitung abzuleiten. Dies soll jedoch auf der Grundlage der Version 2 der Common Guidelines geschehen, die vor kurzem veröffentlicht wurde. Diese wurde deutlich überarbeitet und wurde daher nicht mehr in dieser Arbeit verwendet. Hier werden beim Herausgeber noch einmal die Kennzahlen der Wichtigkeit aus Sicht der Working Group angefragt.

Die Zusammenlegung der Archive stellt aus Sicht des Autors ein wichtiger Schritt für die Zukunft der Erdbeobachtungsdatenarchive der beiden betrachteten Standorte dar. Auch wenn die wirklichen Kosten auf beiden Seiten der Rechnung noch nicht genau fest stehen, so zeigte sich doch ein mögliches großes Einsparungspotenzial bei der benötigten Hardware.

Im letzten Kapitel sollte exemplarisch gezeigt werden, wie eine Datenstruktur in der Product Library zur Speicherung von Dokumenten neben den Daten der selben Mission aufgebaut sein könnte. Hier gibt es noch weiteres Forschungspotenzial zur besten Eingliederung des neuen Produkttyps sowie zu den enthaltenen Metainformationen, damit diese eventuell einerseits standardkonform aber andererseits noch einfach zu pflegen sind.

## Literaturverzeichnis

### Daenzer 2002

DAENZER, W. F. (Hrsg.): *Systems Engineering: Methodik und Praxis*. Bd. 11. Verlag Industrielle Organisation, 2002. – ISBN 3-85743-998-X

### Dech 2006

DECH, Stefan (Hrsg.): *German Remote Sensing Data Center - Status Report 2000-2006*. 2006 1, 3

### Geschonnek 2004

GESCHONNEK, Alexander: *Computer-Forensik: Systemeinbrüche erkennen, ermitteln, aufklären*. dpunkt.verlag, 2004. – ISBN 3-89864-253-4

### Giaretta 2011

GIARETTA, David: *Advanced Digital Preservation*. Springer, 2011. – ISBN 978-3642168086 16

### Kröger u. a. 2009

KRÖGER, Sven ; SCHWINGER, Maximilian ; WEGNER, Max ; WOLFMÜLLER, Meinhard: *Data Handling and Preservation for the TANDEM-X Satellite Mission*. 2009. – <http://elib.dlr.de/60167>

### LTDP 2010

LTDP (Hrsg.): *A Long Term Preservation Framework for Earth Observation Data: LTDP GUIDELINES*. 2010. – DA-09-01c

### LTDP 2012

LTDP (Hrsg.): *Long Term Data Preservation: Earth Observation Preserved Data Set Content*. 2012. – LTDP-GSEG-EOPG-RD-11-0003 17

### Neuroth 2009

NEUROTH, Heike (Hrsg.): *nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*. Bd. 2.0. Hülsbusch, W, 2009. – ISBN 978-3940317483

## A. Anhang

### A.1. Produktmodell MetalItemFile

```

1 <!-- -->
2 <!-- -->
3 <!-- Item Space Configuration IIF -->
4 <!-- -->
5 <!-- -->
6 <!-- Space: DemoDocuSpace -->
7 <!-- Generated by @(#) $Id$, Copyright DLR 2004 -->
8 <!-- Generated on 2013-03-28T13:01:12.437 -->
9
10 <IIF>
11
12 <!-- -->
13 <!-- Complex Item Types -->
14 <!-- -->
15 <item>
16   <administration>
17     <id>dims_op_pl_dfd_XXXXB00000001364475335122</id>
18     <type>ComplexItemType</type>
19     <runtimeType>de.dlr.dfd.dims.pl2.ic.metaItemModel.ComplexItemType</runtimeType>
20   </administration>
21   <specificParameters>
22     <feature key="name" xmlns="http://www.caf.dlr.de/dims/pl/tf">ProductTypeDocumentation</
23       feature>
24     <feature key="base">
25       <feature key="internalName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
26       <feature key="runtimeName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
27       <feature key="representativity" xmlns="http://www.caf.dlr.de/dims/pl/tf">MASTER</feature>
28       <feature key="versionComment" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
29       <feature key="versionCreator" xmlns="http://www.caf.dlr.de/dims/pl/tf">MaxW</feature>
30       <feature key="versionDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
31         :59:35.203</feature>
32       <feature key="version" xmlns="http://www.caf.dlr.de/dims/pl/tf">1.0</feature>
33       <feature key="expirationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
34       <feature key="creationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
35         :59:35.203</feature>
36       <feature key="semantic" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
37       <feature key="space" xmlns="http://www.caf.dlr.de/dims/pl/tf">DemoDocuSpace</feature>
38       <feature key="accessibleSpace" xmlns="http://www.caf.dlr.de/dims/pl/tf">ProductSpace</
39       feature>
40     <feature key="shortMeaning" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
41   </specificParameters>
42 </item>
43 <feature key="isVirtual" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
44 <feature key="isAbstract" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>

```

```

40 <feature key="superType" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_CT_Product_BaseProductSpace</feature>
41 <feature key="isIdentifiedBy" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_XXXB00000001364475228408</feature>
42 <feature key="isIdentifiedBy" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_BA_revision_BaseProductSpace</feature>
43 <feature key="isIdentifiedBy" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_XXXB00000001364475411723</feature>
44 <feature key="isBrieflyDescribedBy" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_XXXB00000001364475228408</feature>
45 <feature key="isBrieflyDescribedBy" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_BA_revision_BaseProductSpace</feature>
46 <feature key="isBrieflyDescribedBy" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_XXXB00000001364475411723</feature>
47 <feature key="isDescribedBy" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_XXXB00000001364475251697</feature>
48 <feature key="isDescribedBy" xmlns="http://www.caf.dlr.de/dims/pl/tf">
    dims_op_pl_dfd_XXXB00000001364475549972</feature>
49 </specificParameters>
50 </item>
51
52 <!-- -->
53 <!-- Structure Item Types -->
54 <!-- -->
55
56 <!-- -->
57 <!-- Structure Attributes -->
58 <!-- -->
59
60 <!-- -->
61 <!-- Basic Attributes -->
62 <!-- -->
63 <item>
64 <administration>
65 <id>dims_op_pl_dfd_XXXB00000001364475228408</id>
66 <type>BasicAttribute</type>
67 <runtimeType>de.dlr.dfd.dims.pl2.ic.metaItemModel.BasicAttribute</runtimeType>
68 </administration>
69 <specificParameters>
70 <feature key="name" xmlns="http://www.caf.dlr.de/dims/pl/tf">productType</feature>
71 <feature key="base">
72 <feature key="internalName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
73 <feature key="runtimeName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
74 <feature key="representativity" xmlns="http://www.caf.dlr.de/dims/pl/tf">MASTER</feature>
75 <feature key="versionComment" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
76 <feature key="versionCreator" xmlns="http://www.caf.dlr.de/dims/pl/tf">MaxW</feature>
77 <feature key="versionDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
    :53:48.404</feature>
78 <feature key="version" xmlns="http://www.caf.dlr.de/dims/pl/tf">1.0</feature>
79 <feature key="expirationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
80 <feature key="creationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
    :53:48.404</feature>
81 <feature key="semantic" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
82 <feature key="space" xmlns="http://www.caf.dlr.de/dims/pl/tf">DemoDocuSpace</feature>
83 <feature key="accessibleSpace" xmlns="http://www.caf.dlr.de/dims/pl/tf">ProductSpace</
    feature>

```

```

84     <feature key="shortMeaning" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
85 </feature>
86 <feature key="isList" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
87 <feature key="multiplicity" xmlns="http://www.caf.dlr.de/dims/pl/tf">1</feature>
88 <feature key="isPresentable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
89 <feature key="isSearchable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
90 <feature key="isUpdateable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
91 <feature key="isNullable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
92 <feature key="onlyValidAllowed" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
93 <feature key="isIdCandidate" xmlns="http://www.caf.dlr.de/dims/pl/tf">true</feature>
94 <feature key="isOfType" xmlns="http://www.caf.dlr.de/dims/pl/tf">BIT_short_string</feature>
95 >
96 </specificParameters>
97 </item>
98 <item>
99 <administration>
100 <id>dims_op_pl_dfd_XXXXB00000001364475251697</id>
101 <type>BasicAttribute</type>
102 <runtimeType>de.dlr.dfd.dims.pl2.ic.metaItemModel.BasicAttribute</runtimeType>
103 </administration>
104 <specificParameters>
105 <feature key="name" xmlns="http://www.caf.dlr.de/dims/pl/tf">productTypeDescription</feature>
106 <feature key="base">
107 <feature key="internalName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
108 <feature key="runtimeName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
109 <feature key="representativity" xmlns="http://www.caf.dlr.de/dims/pl/tf">MASTER</feature>
110 >
111 <feature key="versionComment" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
112 <feature key="versionCreator" xmlns="http://www.caf.dlr.de/dims/pl/tf">MaxW</feature>
113 <feature key="versionDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
114 :54:11.695</feature>
115 <feature key="version" xmlns="http://www.caf.dlr.de/dims/pl/tf">1.0</feature>
116 <feature key="expirationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
117 <feature key="creationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
118 :54:11.695</feature>
119 <feature key="semantic" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
120 <feature key="space" xmlns="http://www.caf.dlr.de/dims/pl/tf">DemoDocuSpace</feature>
121 <feature key="accessibleSpace" xmlns="http://www.caf.dlr.de/dims/pl/tf">ProductSpace</feature>
122 <feature key="shortMeaning" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
123 </feature>
124 <feature key="isList" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
125 <feature key="multiplicity" xmlns="http://www.caf.dlr.de/dims/pl/tf">1</feature>
126 <feature key="isPresentable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
127 <feature key="isSearchable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
128 <feature key="isUpdateable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
129 <feature key="isNullable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
130 <feature key="onlyValidAllowed" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
131 <feature key="isIdCandidate" xmlns="http://www.caf.dlr.de/dims/pl/tf">true</feature>
132 <feature key="isOfType" xmlns="http://www.caf.dlr.de/dims/pl/tf">BIT_std_string</feature>
133 </specificParameters>
134 </item>
135 <item>
136 <administration>
137 <id>dims_op_pl_dfd_XXXXB00000001364475411723</id>
138 <type>BasicAttribute</type>

```

```

135 <runtimeType>de.dlr.dfd.dims.pl2.ic.metaItemModel.BasicAttribute</runtimeType>
136 </administration>
137 <specificParameters>
138 <feature key="name" xmlns="http://www.caf.dlr.de/dims/pl/tf">documentTitle</feature>
139 <feature key="base">
140 <feature key="internalName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
141 <feature key="runtimeName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
142 <feature key="representativity" xmlns="http://www.caf.dlr.de/dims/pl/tf">MASTER</feature>
143 >
144 <feature key="versionComment" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
145 <feature key="versionCreator" xmlns="http://www.caf.dlr.de/dims/pl/tf">MaxW</feature>
146 <feature key="versionDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
147 :56:51.722</feature>
148 <feature key="version" xmlns="http://www.caf.dlr.de/dims/pl/tf">1.0</feature>
149 <feature key="expirationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
150 <feature key="creationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
151 :56:51.722</feature>
152 <feature key="semantic" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
153 <feature key="space" xmlns="http://www.caf.dlr.de/dims/pl/tf">DemoDocuSpace</feature>
154 <feature key="accessibleSpace" xmlns="http://www.caf.dlr.de/dims/pl/tf">ProductSpace</
155 feature>
156 <feature key="shortMeaning" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
157 </feature>
158 <feature key="isList" xmlns="http://www.caf.dlr.de/dims/pl/tf">>false</feature>
159 <feature key="multiplicity" xmlns="http://www.caf.dlr.de/dims/pl/tf">1</feature>
160 <feature key="isPresentable" xmlns="http://www.caf.dlr.de/dims/pl/tf">>false</feature>
161 <feature key="isSearchable" xmlns="http://www.caf.dlr.de/dims/pl/tf">>false</feature>
162 <feature key="isUpdateable" xmlns="http://www.caf.dlr.de/dims/pl/tf">>false</feature>
163 <feature key="isNullable" xmlns="http://www.caf.dlr.de/dims/pl/tf">>false</feature>
164 <feature key="onlyValidAllowed" xmlns="http://www.caf.dlr.de/dims/pl/tf">>false</feature>
165 <feature key="isIdCandidate" xmlns="http://www.caf.dlr.de/dims/pl/tf">>true</feature>
166 <feature key="isOfType" xmlns="http://www.caf.dlr.de/dims/pl/tf">BIT_std_string</feature>
167 </specificParameters>
168 </item>
169 <item>
170 <administration>
171 <id>dims_op_pl_dfd_XXXXB00000001364475549972</id>
172 <type>BasicAttribute</type>
173 <runtimeType>de.dlr.dfd.dims.pl2.ic.metaItemModel.BasicAttribute</runtimeType>
174 </administration>
175 <specificParameters>
176 <feature key="name" xmlns="http://www.caf.dlr.de/dims/pl/tf">documentDate</feature>
177 <feature key="base">
178 <feature key="internalName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
179 <feature key="runtimeName" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
180 <feature key="representativity" xmlns="http://www.caf.dlr.de/dims/pl/tf">MASTER</feature>
181 >
182 <feature key="versionComment" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
183 <feature key="versionCreator" xmlns="http://www.caf.dlr.de/dims/pl/tf">MaxW</feature>
184 <feature key="versionDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
185 :59:09.971</feature>
186 <feature key="version" xmlns="http://www.caf.dlr.de/dims/pl/tf">1.0</feature>
187 <feature key="expirationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
188 <feature key="creationDate" xmlns="http://www.caf.dlr.de/dims/pl/tf">2013-03-28T12
189 :59:09.971</feature>
190 <feature key="semantic" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
191 <feature key="space" xmlns="http://www.caf.dlr.de/dims/pl/tf">DemoDocuSpace</feature>

```

```

185     <feature key="accessibleSpace" xmlns="http://www.caf.dlr.de/dims/pl/tf">ProductSpace</
        feature>
186     <feature key="shortMeaning" xmlns="http://www.caf.dlr.de/dims/pl/tf"></feature>
187 </feature>
188 <feature key="isList" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
189 <feature key="multiplicity" xmlns="http://www.caf.dlr.de/dims/pl/tf">1</feature>
190 <feature key="isPresentable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
191 <feature key="isSearchable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
192 <feature key="isUpdateable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
193 <feature key="isNullable" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
194 <feature key="onlyValidAllowed" xmlns="http://www.caf.dlr.de/dims/pl/tf">false</feature>
195 <feature key="isIdCandidate" xmlns="http://www.caf.dlr.de/dims/pl/tf">true</feature>
196 <feature key="isOfType" xmlns="http://www.caf.dlr.de/dims/pl/tf">BIT_date_f3</feature>
197 </specificParameters>
198 </item>
199
200 <!-- -->
201 <!-- Constants -->
202 <!-- -->
203
204 <!-- -->
205 <!-- Explicit Reference Attributes -->
206 <!-- -->
207
208 <!-- -->
209 <!-- Specific Instances -->
210 <!-- -->
211 </IIF>

```

## **A.2. LTDP Common Guidelines**